



Previsão de Consumo de Energia Elétrica e do Preço da Eletricidade através de Redes Neurais de Regressão Generalizada

José Pedro Ferreira Pelicano Paulos

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Orientador: Professor Doutor José Nuno Fidalgo

14 de Julho de 2017

Resumo

Um dos principais desafios impostos para o século XXI na área da energia, passa por realizar uma gestão coordenada do mercado da eletricidade, permitindo uma cooperação adequada com a produção de energia, permitindo integrar ferramentas computacionais que sustentem um setor elétrico viável e eficaz.

O preço da eletricidade e o consumo energético apresentam-se cada vez mais como os dois grandes pilares do setor elétrico, com a sua análise e previsão a ser uma peça de capital importância para a estável manutenção de um mercado que agora é liberalizado. Com a liberalização do mercado da eletricidade, a previsão de preços e consumo torna-se fundamental para que se desenvolvam tomadas de decisão e para que se estabeleçam estratégias quer por parte dos operadores das redes quer dos operadores de mercado. Ainda que, pela análise do preço da eletricidade, se revelem situações monetariamente estáveis, sobressaem também cenários repletos de irregularidades e daí uma resultante dificuldade em estimar os preços, que podem penalizar o desempenho global da previsão. O consumo de energia, por sua vez, impõe curvas que refletem uma robustez característica, e que permitem realizar previsões razoavelmente satisfatórias.

No âmbito da presente dissertação, pretende-se realizar a previsão de preço da eletricidade e do consumo de energia elétrica, com base em redes neuronais artificiais (*ANN*). Ora, pretende-se verificar o desempenho de diversos tipos de *ANN*, no caso, uma rede neuronal *Feed-Forward* convencional e uma rede neuronal de regressão generalizada (*GRNN*). Tendo como referência valores obtidos pela análise clássica de séries temporais (*Box-Jenkins*), pretende-se comparar os diferentes tipos de redes neuronais quanto, por exemplo, à sua capacidade de resposta a flutuações do mercado, desempenho global e tempos de treino.

Importa, claro, ter em consideração uma análise e seleção das diferentes variáveis a estudar, até porque a representação de um modelo final de previsão deve ter em consideração apenas aquelas que de facto afetam a previsão do preço da eletricidade ou do consumo.

A abordagem pelas *GRNN* é uma solução já com alguns anos, mas tem ressurgido recentemente realçando um bom desempenho, rapidez da fase de treino e facilidade de parametrização. A integração deste modelo em estudos comparativos com métodos mais convencionais permite obter soluções robustas com tempos de treino bastante reduzidos, comparativamente às *ANN*. Percebe-se que, de uma forma geral, o desempenho das *GRNN*, ainda que levemente inferior ao das *ANN*, é excelente, mas associado a tempos de treino muito inferiores.

Palavras-chave: Indicadores de desempenho, previsão do consumo de energia elétrica, previsão do preço da eletricidade, redes neuronais, redes neuronais de regressão generalizada.

Abstract

One of the main challenges of the 21st century in the energy sector is to achieve a coordinated management of the electricity market, enabling an adequate cooperation with energy production systems and allowing the integration of computational tools in order to achieve a viable and efficient electricity sector.

Over time, the electricity price and energy consumption are increasingly growing their weight as the main foundations of the electrical sector, with their analysis and forecasts being targeted as key elements for the stable maintenance of a market that has only been liberalized for a short time. Hence, load and electricity price forecasting are becoming crucial tools at forging the basics steps to make institutional decisions by both network and market operators. Regardless of the stable monetary current status, there are scenarios where electricity price inquiries can suffer from irregularities and therefore a difficulty in estimating their prices, which may inflict a handicap on the overall performance of the forecasts. The energy consumption, in turn, introduces curves that reflect a characteristic robustness and thus, reasonably satisfactory forecasts results.

In this dissertation, it is intended to carry out a research about the load and electricity price forecast through artificial neural networks (ANN). The aim of this drill is to check the performance of several ANN strains, for instance, a conventional Feed-Forward neural network and a Generalized Regression Neural Network (GRNN). Based on the results obtained by the classical time-series analysis (Box-Jenkins), we expect to compare the different types of neural networks, such as their ability to acknowledge market fluctuations, overall performance and training times.

It is intended, by all means, to take into account an analysis and draft of the specific variables to be selected, since the representation of a final model of forecasting should only consider those whose, in fact, affect the load and electricity price forecasts.

The GRNN approach is a solution dated from 1991, but has recently re-emerged emphasizing good performance, fast runtimes and ease of parameterization. The merging of this model in comparative studies with more conventional methods allows us to obtain more sturdy solutions with shortened runtimes, when compared to ANN. Overall, the performance of the GRNN, although slightly inferior to that of the ANN, is excellent, but linked to much lower training times.

Keywords: *Load forecasting, electricity price forecasting, neural networks, generalized regression neural networks, performance indicators.*

Agradecimentos

Termina mais uma etapa da minha formação. Não posso deixar de manifestar os meus sinceros agradecimentos às pessoas que de uma forma ou de outra participam na sua concretização. Não sou um sentimental, prefiro um silêncio soturno a um banquete de celebração rejubilante. Ainda assim, a cooperação destas pessoas desvenda um trajeto inicialmente agitado mas sempre abrigado por predicados que abraçam uma crónica finalmente triunfante. Partidário de análises quanto ao primarismo das primeiras etapas, o culminar deste roteiro é evidentemente nutrido pela influência dessas pessoas.

Em primeiro, agradeço aos meus pais e irmãos, pela consciência da convicção, pelas oportunidades, motivação e pela força que se demonstraram como essenciais ao longo de todo o meu processo académico.

Agradeço a todos os meus familiares pelo ânimo e incentivo que permitiram exceder as duras barreiras que fui enfrentando.

Às grandes amizades conseguidas, pela motivação e pelo companheirismo que sempre demonstraram.

Por último, agradeço ao meu orientador, Professor Doutor José Nuno Fidalgo a sua disponibilidade constante, e generosa partilha de conhecimento que se demonstraram como precípuas para a realização desta dissertação.

José Pedro Ferreira Pelicano Paulos

“Life can only be understood backwards, but it must be lived forwards.”

Søren Kierkegaard

Conteúdo

1	Introdução	1
1.1	Enquadramento e Motivação	1
1.2	Objetivos	2
1.3	Informação e Caso de Estudo	3
1.4	Estrutura da Dissertação	3
2	Revisão Bibliográfica	5
2.1	Introdução	5
2.2	Modelos Matemáticos e Computacionais	5
2.3	Previsão	6
2.3.1	Métodos de Previsão	8
2.4	Previsão pela Análise de Séries Temporais - <i>Box-Jenkins</i>	9
2.5	Redes Neurais Artificiais	13
2.5.1	História	13
2.5.2	Funcionamento de uma Rede Neuronal	14
2.5.3	Arquitetura de uma Rede Neuronal	17
2.5.4	Aprendizagem de uma Rede Neuronal	22
2.5.5	Algoritmo de Treino <i>Levenberg-Marquardt</i>	24
2.5.6	Aplicações de Redes Neurais	26
2.6	O Setor Elétrico em Portugal	28
2.6.1	Preço da Eletricidade	28
2.6.2	Consumo de Energia Elétrica	30
3	Ferramenta de previsão e procedimentos utilizados	33
3.1	Rede Neuronal de Regressão Generalizada	33
3.1.1	Introdução	33
3.1.2	Abordagem e Estrutura	33
3.1.3	Normalização de Dados e Parâmetro de Suavização	36
3.2	Obtenção e Tratamento de Informação	37
3.2.1	Pré-processamento de Dados	38
3.2.2	Ferriados	39
3.2.3	Indicadores de Desempenho	40
4	Metodologia Utilizada	43
4.1	Introdução	43
4.2	<i>Software</i> Utilizado	43
4.2.1	Organização e Tratamento de Dados	43
4.2.2	Execução de Simulações e Cálculo de Indicadores de Performance	43

4.2.3	Ferramenta para Criação e Treino de <i>GRNN</i>	44
4.2.4	Ferramenta para Criação e Treino de <i>Feed-Forward ANN</i>	44
4.3	Aplicação Desenvolvida para Realização de Simulações	44
4.3.1	Pré-processamento de Dados	46
4.3.2	Amostragem de Dados	46
4.3.3	Previsão através de <i>GRNN</i>	52
4.3.4	Previsão através de <i>Feed-Forward ANN</i>	52
4.3.5	Apresentação de Resultados	53
5	Apresentação e Análise de Resultados	55
5.1	Escolha do Melhor Modelo para Previsões	55
5.1.1	Preço da Eletricidade	55
5.1.2	Consumo de Energia Elétrica	71
5.2	Previsões Finais	87
5.2.1	Previsões Finais do Preço da Eletricidade	88
5.2.2	Previsões Finais do Consumo de Energia Elétrica	91
5.2.3	Previsão Experimental do Preço da Eletricidade com Seleção Horária	93
6	Conclusões e Trabalho Futuro	95
6.1	Conclusões	95
6.1.1	1ª Fase: Pré-processamento de Dados	95
6.1.2	2ª Fase: Análise do Comportamento das Séries	95
6.1.3	3ª Fase: Estudo e Escolha dos Métodos de Amostragem de Dados	96
6.1.4	4ª Fase: Identificação dos Modelos Finais para Previsão	96
6.1.5	5ª Fase: Realização das Previsões	96
6.1.6	Considerações Finais	97
6.2	Aplicações Práticas da Previsão de Consumo e Preço	97
6.3	Dificuldades Encontradas	97
6.4	Satisfação de Objetivos e Trabalhos Futuros	98
	Referências	99

Lista de Figuras

2.1	Diagrama de carga da RNT a 10 de dezembro de 2012	9
2.2	Procedimento para construção de um modelo <i>ARIMA</i>	11
2.3	Exemplo de funções de autocorrelação e autocorrelação parcial.	12
2.4	Estrutura geral de um neurónio.	17
2.5	Exemplo da função sigmoide para as inclinações de $k = 0.5$, $k = 1.0$ e $k = 2.0$. .	19
2.6	Diferenças entre os modos de ligação entre neurónios.	20
2.7	Rede neuronal <i>Feed-Forward</i> de apenas uma camada.	20
2.8	Rede neuronal <i>Feed-Forward</i> de múltiplas camadas.	21
2.9	Preço médio anual do mercado diário desde 2015 (€/MWh).	29
2.10	Preço médio do mercado diário de 2016 (€/MWh).	29
2.11	Estatística diária do SEE a 26 de abril de 2017	30
2.12	Estatística diária do consumo elétrico do ano de 2016 (MWh).	31
3.1	Estrutura geral de uma <i>GRNN</i>	34
3.2	Arquitetura de uma <i>GRNN</i>	35
3.3	Algoritmo de funcionamento de uma <i>GRNN</i>	37
4.1	Interface gráfica da aplicação desenvolvida.	45
4.2	Exemplo de uma amostragem aleatória.	47
4.3	Exemplo de uma amostragem ordenada.	48
4.4	Exemplo de uma amostragem de semanas misturadas.	49
4.5	Exemplo de uma amostragem adaptativa hora-a-hora.	50
4.6	Exemplo de uma amostragem adaptativa dia-a-dia.	51
5.1	Evolução do preço da eletricidade em 2016 (€/MWh).	56
5.2	Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2016.	56
5.3	Evolução do preço médio semanal da eletricidade nos anos de 2011 a 2016 (€/MWh) .	57
5.4	Evolução do preço da eletricidade em 2011 (€/MWh).	58
5.5	Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2011.	58
5.6	Evolução do preço da eletricidade em 2012 (€/MWh).	59
5.7	Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2012.	59
5.8	Evolução do preço da eletricidade em 2013 (€/MWh).	60
5.9	Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2013.	60
5.10	Evolução do preço da eletricidade em 2014 (€/MWh).	61

5.11 Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2014.	61
5.12 Evolução do preço da eletricidade em 2015 (€/MWh).	62
5.13 Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2014.	62
5.14 Evolução do consumo de energia elétrica em 2016 (MWh).	72
5.15 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2016.	72
5.16 Evolução do consumo médio semanal de energia elétrica e nos anos de 2011 a 2016 (MWh)	73
5.17 Evolução do consumo de energia elétrica em 2011 (MWh).	74
5.18 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2011.	74
5.19 Evolução do consumo de energia elétrica em 2012 (MWh).	75
5.20 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2012.	75
5.21 Evolução do consumo de energia elétrica em 2013 (MWh).	76
5.22 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2013.	76
5.23 Evolução do consumo de energia elétrica em 2014 (MWh).	77
5.24 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2014.	77
5.25 Evolução do consumo de energia elétrica em 2015 (MWh).	78
5.26 Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2015.	78
5.27 Resultados finais da previsão do preço em 2016 via ANN, por seleção autoadaptativa.	90
5.28 Resultados finais da previsão do preço em 2016 via GRNN, por seleção autoadaptativa.	90
5.29 Resultados finais da previsão do consumo em 2016 via ANN, por seleção autoadaptativa.	93
5.30 Resultados finais da previsão do consumo em 2016 via GRNN, por seleção autoadaptativa.	93

Lista de Tabelas

2.1	Tipo de modelo <i>Box-Jenkins</i> a utilizar para diferentes comportamentos de ACF e PACF.	12
2.2	Tipos de funções de ativação	18
5.1	Resultados de previsão do preço da eletricidade de 2011 a 2016, via <i>Box-Jenkins</i>	63
5.2	Listagem de modelos a estudar para a previsão do preço da eletricidade	64
5.3	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2016.	65
5.4	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2016.	65
5.5	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2015.	66
5.6	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2015.	66
5.7	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2014.	67
5.8	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2014.	67
5.9	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2013.	68
5.10	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2013.	68
5.11	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2012.	69
5.12	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2012.	69
5.13	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2011.	70
5.14	Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2011.	70
5.15	Resumo dos modelos a utilizar para a previsão do preço da eletricidade, por ano.	71
5.16	Resultados da previsão do consumo de energia elétrica de 2011 a 2016, via <i>Box-Jenkins</i>	79
5.17	Listagem de modelos a estudar para a previsão do consumo de energia elétrica . .	80
5.18	Resultados dos testes de previsão do consumo de energia elétrica, por redes neuronais, para os modelos Base e A, relativos ao ano de 2016.	81
5.19	Resultados dos testes de previsão do consumo de energia elétrica, por redes neuronais, para os modelos B e C, relativos ao ano de 2016.	81

5.20	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2015.	82
5.21	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2015.	82
5.22	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2014.	83
5.23	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2014.	83
5.24	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2013.	84
5.25	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2013.	84
5.26	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2012.	85
5.27	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2012.	85
5.28	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2011.	86
5.29	Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2011.	86
5.30	Resumo dos modelos a utilizar para a previsão do consumo de energia elétrica, por ano.	87
5.31	Resultados comparativos da previsão do preço da eletricidade, entre métodos de seleção ordenado e adaptativo, para os anos de 2011 a 2016.	88
5.32	Resultados finais e comparativos da previsão do preço da eletricidade, entre <i>GRNN</i> e <i>ANN</i> , para os anos de 2011 a 2016.	89
5.33	Resultados comparativos da previsão do consumo de energia elétrica, entre métodos de seleção ordenado e adaptativo, para os anos de 2011 a 2016.	91
5.34	Resultados finais e comparativos da previsão do consumo de energia elétrica, entre <i>GRNN</i> e <i>ANN</i> , para os anos de 2011 a 2016.	92
5.35	Resultados comparativos da previsão experimental do preço da eletricidade, entre <i>GRNN</i> e <i>ANN</i> , por seleção adaptativa horária, para o ano de 2016.	94

Abreviaturas e Símbolos

ACF	Autocorrelation Function
ANN	Artificial Neural Network
AR	Autoregressive
ARIMA	Autoregressive-Integrated-Moving-Average
ARMA	Autoregressive-Moving-Average
DDS	Dia da Semana
GNA	Gauss-Newton Algorithm
GRNN	Generalized Regression Neural Network
LM	Levenberg-Marquardt
LMA	Levenberg-Marquardt Algorithm
MA	Moving Average
MAE	Mean Average Error
MAPE	Mean Absolute Percentage Error
MIBEL	Mercado Ibérico de Eletricidade
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
OMIE	Operador do Mercado Ibérico
PACF	Partial Autocorrelation Function
PRE	Produção em Regime Especial
PRO	Produção em Regime Ordinário
REN	Redes Elétricas Nacionais
RMSE	Root Mean Squared Error
SEE	Sistema Elétrico de Energia
UE	União Europeia

Lista de Símbolos

\cos	Função cosseno
d	Parâmetro I
e	Erro de previsão
H	Matriz <i>Hessiana</i>
h	Horas
J	Matriz Jacobiana
MW	Megawatt
MWh	Megawatt-hora
m	Minutos
mod	Função modular
p	Parâmetro AR
q	Parâmetro MA
s	Segundos
sign	Função sinal
sen	Função seno
T	Conjunto de dados total
T_{test}	Conjunto de treino
T_{train}	Conjunto de teste
\tanh	Função tangente hiperbólica
W	Watt
W_{test}	Peso do conjunto de teste
W_{train}	Peso do conjunto de treino
w_i	Peso da rede neuronal para o neurónio i
x	Input da rede neuronal
y	Output da rede neuronal
€	Euro
Σ	Somatório
$\nabla \xi$	Gradiente da função de custo
ξ	Função de custo
η	Taxa de aprendizagem
σ	Parâmetro de suavização

Capítulo 1

Introdução

A presente dissertação foi realizada no âmbito do Mestrado Integrado em Engenharia Electrotécnica e de Computadores (MIEEC), da Faculdade de Engenharia da Universidade do Porto (FEUP).

Os dados utilizados ao longo de todo o processo foram fornecidos ou transferidos das plataformas de resultados da REN (Redes Energéticas Nacionais) e da OMIE (Operador do Mercado Ibérico).

Para efeito de tratamento e análise de dados e para as respetivas previsões foi utilizado o *software Microsoft® Excel*, o *IBM SPSS Statistics*, e o *MathWorks MATLAB*. A escrita do documento foi realizada em sistema \LaTeX .

O presente capítulo introdutório encontra-se dividido em quatro secções, nomeadamente:

- Enquadramento e motivação face ao tema em estudo nesta dissertação;
- Objetivos definidos;
- Caso de estudo;
- Estrutura da dissertação

1.1 Enquadramento e Motivação

A revolução tecnológica, científica, económica e social que tem vindo a caracterizar a evolução das últimas décadas é essencial para o aumento da qualidade de vida de todos os cidadãos.

Associado a esta evolução, tem-se verificado um crescente aumento da dependência energética, identificada por uma exploração radical de combustíveis fósseis que, desde sempre, foram um dos pilares, de uma exploração descontrolada graças à sua abundância e ao seu elevado aproveitamento energético.

Perante a dependência mundial de fontes de energia convencionais e considerando os seus efeitos ambientais, torna-se imperativo o desenvolvimento e exploração de novas fontes de energia, aptas para possibilitar uma sustentabilidade energética eficaz, acompanhada por uma substancial redução da pegada ambiental.

Os mercados de eletricidade, agora sob a forma de bolsa, são baseados nas curvas de procura e oferta e são caracterizados por uma forte volatilidade dos preços finais, que podem variar sobre dezenas de euros em intervalos de uma hora. Desta forma, o risco do mercado é considerado como bastante elevado pelo que, para os diversos intervenientes, se torna essencial possuírem previsões do preço da eletricidade que proporcionem uma melhor avaliação do risco das suas propostas de compra ou venda.

A eletricidade, enquanto mercadoria, apresenta propriedades únicas tais como a impossibilidade de ser armazenada de forma eficiente e economicamente viável e, daí, a produção ter que satisfazer de forma constante e instantânea as necessidades do diagrama de carga é um dos principais responsáveis pela incerteza associada ao preço da eletricidade. Para além do consumo, existem outras peças fundamentais para a atribuição do preço da eletricidade. Note-se a crescente aposta em fontes de energia renováveis, que se têm vindo a destacar com elevadas capacidades de produção resultando numa razoável redução no preço da eletricidade.

Assim, a estratificação do modelo atual do setor elétrico permite que, com os dados de previsão do consumo de energia elétrica, se consiga realizar uma gestão coordenada das produções, garantindo um aditamento regulado da produção renovável.

1.2 Objetivos

Na presente dissertação é definido como objetivo geral a comparação de desempenho e tempos de execução entre redes neuronais de regressão generalizada e redes neuronais convencionais *feed-forward*. Querendo garantir a execução deste exercício, será necessário:

- Obtenção de dados relativos à produção e consumo de energia elétrica e do preço da eletricidade para os anos em estudo;
- Criação de uma ferramenta que possibilite a automatização de simulação e pré-processamento de dados;
- Desenvolvimento de diversas técnicas de seleção de dados;
- Escolha dos melhores modelos para cada variável a prever;
- Desenvolvimento das previsões para ambos os métodos, para cada ano em estudo e para diversas técnicas de seleção de dados;
- Avaliação dos indicadores de desempenho;

1.3 Informação e Caso de Estudo

Os dados utilizados para construção e validação dos modelos de previsão desenvolvidos foram fornecidos ou transferidos das plataformas de resultados da REN (Redes Energéticas Nacionais) e da OMIE (Operador do Mercado Ibérico). No caso, faz-se uso dos valores de produção de energia elétrica (por tipo de geração), consumo de energia elétrica e preço da eletricidade, para os anos de 2011 a 2016, inclusive.

1.4 Estrutura da Dissertação

Para além do capítulo introdutório, a presente dissertação dispõe de mais cinco capítulos. No Capítulo 2 apresenta-se o estado da arte relativo a uma revisão teórica de métodos de previsão, redes neuronais e também uma breve referência do setor elétrico em Portugal. No Capítulo 3, clarifica-se aquelas metodologias e técnicas utilizadas para obtenção e tratamento de dados. Já no Capítulo 4 são feitas referências ao *software* utilizado para organização e pré-processamento. Para além disso, são demonstradas as funcionalidades da aplicação desenvolvida para a realização da seleção e previsão de dados. No Capítulo 5 é completada uma apresentação e análise dos resultados obtidos. Finalmente, o Capítulo 6 serve como consolidação e síntese das conclusões obtidas no decurso de toda a dissertação.

Capítulo 2

Revisão Bibliográfica

2.1 Introdução

Neste capítulo é elaborada uma síntese de conceitos associados à previsão, os diferentes métodos de previsão, do desenvolvimento de modelos associados e ainda uma caracterização do comportamento do preço da eletricidade e da variação da carga.

2.2 Modelos Matemáticos e Computacionais

O estudo científico da realidade, ou de uma determinada situação sob observação, centra-se sempre em torno de um modelo. Um modelo é uma representação simplificada da realidade, uma interpretação fragmentada de uma dada ocorrência, e que se apresenta apenas como uma única visão num vasto leque de possíveis cenários. Ora, garante-se que, para o estudo de um dado fenómeno complexo, possam ser estabelecidos vários modelos que poderão ser úteis para compreender os efeitos dos seus componentes e até prever comportamentos. Um modelo, ainda que insuficiente ou inconsistente, pode ser usado caso não seja possível obter uma melhor solução.

Ainda assim, neste caso, importa destacar os modelos matemáticos, isto é, modelos que utilizam conceitos e ideais matemáticas para explorar a realidade. Habitualmente desenvolvidos num panorama teórico, os modelos matemáticos são interessantes pela possível manipulação de teoremas já existentes e pela possibilidade de executar cálculos complexos através de aplicações computacionais. Ainda mais, é possível retirar conclusões distintas, sobretudo sustentadas em eventos repetíveis, de forma a garantir a evolução dum modelo, garantindo novas medidas experimentais e o desenvolvimento de avanços e novas teorias. Para além disso, a falta de consenso e de concordância entre ideias e conclusões, pode precipitar a ocorrência de novos estudos e daí a exploração de novas versões dum dado modelo. Facilmente se entende que, para atingir uma meta comum, podem existir diferentes formas de o conseguir.

Habitualmente, no que toca a um modelo, segundo [1], há que ter em consideração três tipos de grandezas que poderão potencialmente integrar o modelo:

1. Aquelas cujos efeitos são negligenciáveis;
2. Aquelas que afetam o modelo mas cujo comportamento resultante não é objeto de estudo;
3. Aquelas cujo comportamento é o alvo de estudo do modelo.

As primeiras, cuja importância é nula, são ignoradas no modelo. As segundas, habitualmente funções, constantes, entre outras, são denominadas *variáveis exógenas*, também chamadas de variáveis independentes, de entradas ou na nomenclatura inglesa, de *inputs*. Os *inputs* podem ser variáveis exógenas mas também podem tomar valores anteriores da variável a prever, ou até dados de calendário. As últimas, aquelas que o modelo pretende estudar, são denominadas *variáveis endógenas*, também conhecidas por variáveis dependentes, saídas, ou em inglês, *outputs*.

A estruturação destes três grupos é crucial para o desenvolvimento de um bom modelo. Nos dois primeiros, caso se erre na escolha de variáveis ou até caso se despreze a quantidade de dados a trabalhar, o modelo pode-se tornar errático ou impreciso. O processo da escolha das variáveis dependentes (i.e., os *outputs*), é também ele essencial e deve-se procurar estudar apenas aquilo que possa ser explicado e que tenha relação com as anteriores. A interligação entre variáveis é o cerne do modelo, e deve ser respeitada de forma a garantir precisão nas conclusões [1].

De forma a otimizar a execução de dado modelo complexo, recorre-se ao conceito de modelo computacional, isto é, um modelo matemático que faz uso de certas valências computacionais, através da simulação e modelagem computacional. Nestes casos, os modelos e o procedimento em questão costumam ser de uma complexidade tal que uma solução analítica seria impraticável. Ora, contrariamente aos modelos analíticos que recorrem à derivação matemática, aqui, a experimentação é baseada num processo de atribuição e ajuste de certos parâmetros do sistema, e pelo estudo das saídas das diversas experiências. Como exemplo deste tipo de modelos, temos os modelos de redes neuronais.

2.3 Previsão

A previsão é um processo que permite antever e analisar a ocorrência de valores futuros, baseados em dados passados, paralelamente com um estudo de tendências. O seu estudo é fundamental para estimar o futuro, sendo uma ferramenta oportuna em diversas aplicações.

Dependendo do tipo de material a prever, o processo pode ser feito com anos de avanço, ou até umas horas ou minutos antes do fenómeno. Quaisquer que sejam as circunstâncias ou os horizontes temporais, uma previsão correta serve de pilar para enquadrar uma boa decisão ou até um planeamento mais completo e eficaz. Ora, enquanto que um planeamento simples e sem o recurso à previsão, para uma determinada situação, se poderá basear em ideias inconsistentes, o enquadramento de uma estratégia com a previsão de dados fornece claramente um planeamento mais eficaz e consistente, garantindo algum nível de solidez quando em comparação com o anterior.

Existem claro, grandezas mais facilmente previstas do que outras, e.g, prever a hora do pôr do sol será certamente mais simples do que prever o sorteio de uma lotaria, até porque existem fatores

aleatórios associados ao último. Portanto, de acordo com [1], a previsibilidade de um determinado evento depende claro de variados parâmetros:

1. Quão bem conhecemos os fatores que contribuem para sua ocorrência;
2. Quanta informação está disponível;
3. De que forma é que estar a realizar uma previsão pode, ou não, interferir no desfecho normal daquilo que queremos prever.

Ora, caso queiramos realizar uma previsão da carga elétrica num dado dia, sabemos que temos praticamente a 100% os três parâmetros respeitados, até porque é um evento minimamente estável e que depende habitualmente das mesmas informações para ser eficazmente previsto. Contudo, se falarmos da previsão da cotação de uma dada moeda, a própria previsão realizada pode interferir no seu valor final.

Uma boa ideia a reter é tentar perceber se um certo evento é facilmente previsível ou não, até porque uma boa previsão é garantida quando se apura que existem evidentes relações de dados históricos e padrões de evolução da informação a estimar. O processamento da informação também é crucial, isto é, perante um conjunto de informação inicial, os dados a relacionar devem ser relativos a um mesmo intervalo temporal e correlacionados com os anteriores, de forma a impedir imperfeições no processo de previsão. Devem-se organizar os recursos e obter informação o tão correta quanto possível, de forma a que a questão seja solucionada com o menor risco possível. A coerência e a relação entre entradas e saídas também deve ser analisada, até porque a ocorrência e o aumento de erro, devido a existência de informação sem qualquer relação, é evidente em certas ocasiões. Não é necessário obter o maior, mas sim o melhor conjunto de informação, de forma a que não seja introduzido “ruído” no processo de previsão.

Além da relevância dos dados a usar, importa também notar que a quantidade de informação necessária deve ser tida em conta. Será com certeza imprudente realizar uma previsão de um conjunto de dados anual, quando apenas se tem disponível informação de certos dias aleatórios de uma semana. Certamente o resultado da previsão será enviesado e a previsão não será muito vantajosa.

Outro critério que contribui para que se obtenha um bom modelo de previsão, é o facto da informação ser representativa do universo de casos a prever, sendo também capaz de filtrar flutuações anómalas e singulares, que tenham ocorrido no passado.

Outra característica a referir é volatilidade destes sistemas, ou seja, são sistemas adaptativos à mudança. Em qualquer ambiente onde se façam testes está presente a possibilidade de variação ou de mudança, mas os sistemas preditivos assumem que o próprio ambiente a prever também é volátil e como tal devem estar preparados para responder a alterações. Ora, se ocorreram mudanças no passado, também poderão ocorrer no futuro. Percebe-se assim que, um modelo de previsão está preparado para julgar tendências, compreender o rumo da informação e não só o seu estado atual.

Estas metodologias de previsão apresentam vastas soluções, consoante o horizonte temporal, tipo de informação, pela suas tendências, etc.. Para além disso, passam por modelos simplistas

e vulgares, mas também por protótipos complexos, como é o caso das redes neuronais. Também podem surgir problemas onde não existe histórico relevante do ambiente a trabalhar, como é o caso da previsão de vendas de um produto novo que nunca foi comercializado, não sendo possível utilizar redes neuronais ou outros métodos baseados em exemplos.

Como referido, a previsão nada mais é do que uma ferramenta essencialmente estatística que permite auxiliar na tomada de decisões e que pode funcionar como uma forma de coadjuvar o planeamento estratégico para diversos cenários. Pretendendo estimar o futuro da forma mais precisa possível, e presente alguma informação, esta deve estar diretamente conectada a um planeamento organizado e realista.

A gestão de informação e as decisões são parcelas de capital importância numa qualquer empresa e, como tal, podem necessitar de recorrer a uma ferramenta deste género, para a mais distintas e específicas situações. Ora, dependendo do ambiente no qual se pretende trabalhar, uma previsão pode ser de três diferentes tipos:

- **Previsão a curto prazo:** necessária habitualmente para organização de pessoal, estimação de produção, procura, ou outros (e.g., pré-despacho ou a operação de compra e venda de energia no mercado da eletricidade; previsões de uma hora até uma semana).
- **Previsão a médio prazo:** para determinação dos requisitos futuros de forma a ser possível contratar pessoal ou comprar equipamento (e.g., previsões mensais);
- **Previsão a longo prazo:** usadas normalmente para planeamento estratégico e de forma a ter em consideração oportunidades de mercado (e.g., previsões semestrais ou anuais).

2.3.1 Métodos de Previsão

Relativamente a métodos de previsão, importa categoriza-los em duas grandes categorias, os métodos qualitativos ou quantitativos. Os primeiros destacam-se pela sua subjetividade, representando técnicas baseadas nas opiniões e juízos de consumidores ou peritos e são particularmente usadas quando a informação passada não existe. Já os métodos quantitativos, são aqueles que têm como grande pilar informação numérica acerca do passado e onde é seguro assumir que certos padrões vão continuar a surgir no futuro.

Como descrito em [2], os métodos quantitativos são aqueles que são baseados em modelos matemáticos ou estatísticos sendo até classificados como determinísticos ou probabilísticos. Nestes, a relação entre uma variável de interesse e as variáveis de entrada é dada por:

$$Y = f(X_1, \dots, X_p) \quad (2.1)$$

Ainda acerca destes, existem variadíssimos métodos ou modelos desenvolvidos certamente tendo em conta o ambiente para o qual seriam usados. Cada modelo é representativo do tipo

de problema a solucionar e pode apresentar características e uma estrutura próprias do evento a prever. A maioria dos métodos quantitativos tem em conta séries temporais, isto é, é seleccionada informação sequencial de um dado intervalo de tempo.

2.4 Previsão pela Análise de Séries Temporais - Box-Jenkins

A análise e previsão de dados através de modelos baseados em séries temporais apresentam-se como um método suficientemente capaz para ser útil nas áreas da indústria e da economia, por exemplo. Como base de algumas das metodologias usadas, encontram-se as séries temporais. Ora, estas previsões, baseiam-se em informação colecionada acerca de uma ou mais séries temporais. Uma série temporal nada mais é do que um conjunto de observações sequenciais no tempo, acerca de um dado parâmetro ou evento [3], uma observação de uma ou mais variáveis ao longo do tempo [4] onde se verifica de algum modo a repetição de um padrão base. Neste caso, a série flutua em torno dessa padrão. Noutros casos, pode não ser possível identificar um padrão.

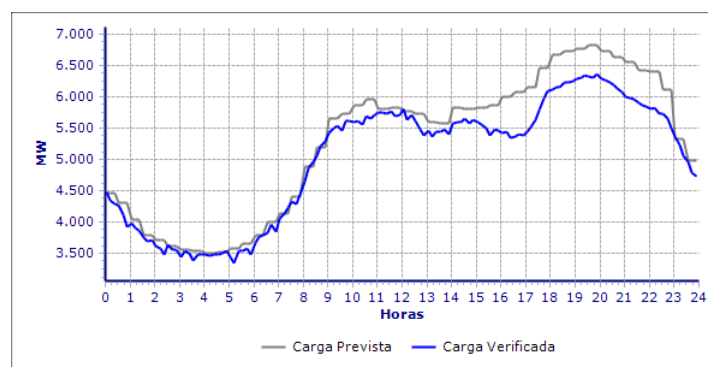


Figura 2.1: Diagrama de carga da RNT a 10 de Dezembro de 2012 [5].

Tal como representado na Figura 2.1, num fragmento da sua totalidade, uma série temporal é então caracterizada por um conjunto de valores de uma grandeza de um dado intervalo temporal, no caso, a variação da carga ao longo de um dia. No caso da carga, é previsível que esta respeite um padrão base, e como tal, o modelo prevê que a curva respeite, de certa forma, esta variação.

Dentro dos modelos mais clássicos para previsão pela análise de séries temporais, temos o modelo *autoregressive integrated moving average* (ARIMA), que se trata de uma generalização do modelo *autoregressive moving average* (ARMA) sendo que ambos se podem tratar de modelos autoregressivos (AR) ou de média móvel (MA).

Os modelos AR foram inicialmente introduzidos por Yule em 1926 [6] e sustentados por Slutsky que em 1937 [7] apresentou os modelos de média móvel (MA). Porém, foi apenas em 1938, por Wold [8] que se combinaram os modelos AR e MA, resultando num modelo ARMA, que possibilitava que se estimassem valores temporais desde que fossem respeitados os parâmetros de ambos os modelos, combinados num só. Isto é, desde que p , o número de termos da parcela AR e de

q fossem especificadas e respeitadas [9]. Então, a parcela *AR* mostra que a evolução da variável em estudo pode ser estimada através da sua própria regressão (i.e. pelos valores passados). Por outro lado, a parcela *MA* indica que a regressão pode ser baseada numa combinação linear entre os diversos erros que ocorreram no passado.

Ora, percebe-se que uma qualquer série x pode ser modelada por uma combinação dos valores passados x_t e/ou dos erros passados e_t :

$$x_t = \phi_1 \cdot x_{t-1} + \phi_2 \cdot x_{t-2} + \cdots + \phi_p \cdot x_{t-p} + e_t - \theta_1 \cdot e_{t-1} - \theta_2 \cdot e_{t-2} - \cdots - \theta_q \cdot e_{t-q} \quad (2.2)$$

Ainda que já implementável, esta solução sugerida por *Wold* [8] (ver expressão 2.2), não foi utilizada até aos anos 60, momento no qual o poder computacional começou a ser suficiente para realizar os cálculos requeridos para otimizar os parâmetros de 2.2, permitindo até a sua utilização no setor económico [9].

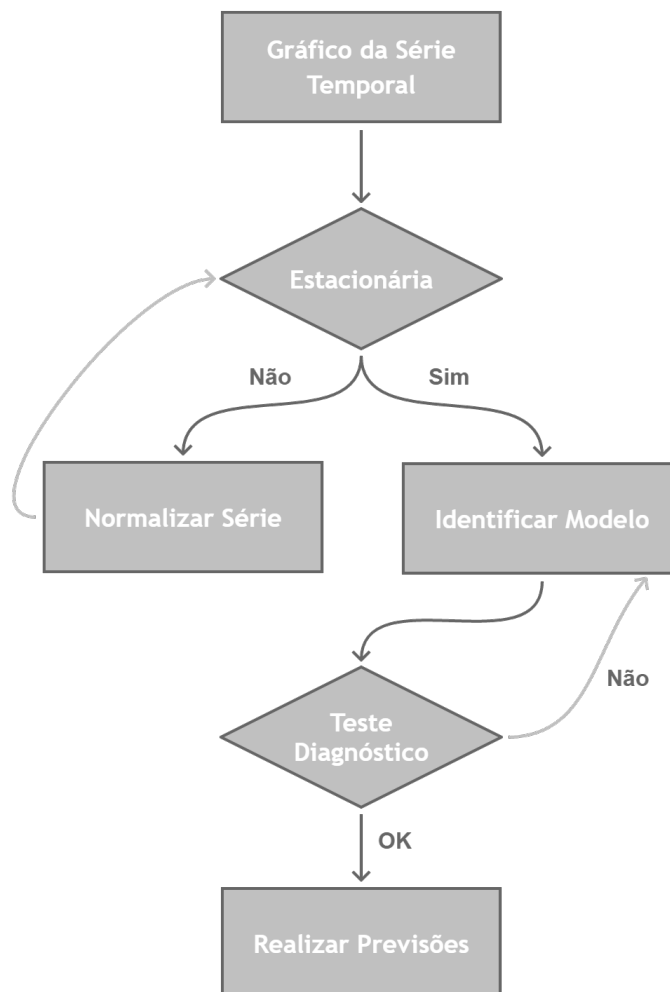
Já em 1976 (ou em 1970 com a edição original) [10], *Box* e *Jenkins* introduziram algumas alterações aos modelos já existentes sugerindo a execução de diversos procedimentos:

- Normalizar a série de dados através da respetiva média e variância;
- Sugere-se o uso de coeficientes de autocorrelação e autocorrelação parcial para determinação eficaz dos parâmetros p e q ;
- Devem-se providenciar aplicações computacionais capazes de identificar valores apropriados para p e q e estimar os parâmetros envolvidos;
- Após estimados os parâmetros do modelo em estudo, propõe-se que se diagnostique o modelo, de forma a verificar se este é, ou não, final.

Esta abordagem é então a metodologia *Box-Jenkins* associada a modelos *ARIMA* onde a letra *I* entre *AR* e *MA* representa a palavra "*Integrated*". Tornando-se academicamente popular nos anos 70, esta metodologia apresentou-se também em diversas investigações (e.g. [11] [12]) onde foi possível superar largamente métodos econométricos clássicos complexos, populares na altura, em diversas situações [9].

Note-se ainda que se consegue identificar o tipo de modelo a utilizar pela verificação dos parâmetros p , d e q . Isto é, um momento *ARIMA* é geralmente classificado como *ARIMA*(p,d,q) onde estas três representam as ordens das parcelas autoregressivas, integradas e de média móvel, respetivamente. Assim, quando dois destes três termos são representados por zeros, estamos perante uma solução baseada num só modelo base (i.e. *ARIMA*(1,0,0) representa um *AR*(1) e assim sucessivamente).

Acerca do procedimento a tomar para construção de um modelo *ARIMA*, deve-se obter o gráfico da série temporal a estudar, verificar se esta é ou não estacionária, identificar o modelo, realizar o diagnóstico e por fim obter as previsões requeridas, tal como indicado na Figura 2.2 baseado em [9] e [13].

Figura 2.2: Procedimento para construção de um modelo *ARIMA*

De forma a identificar o tipo de modelo a utilizar, interessa determinar p e q . Portanto, deve-se fazer uso de conceitos da autocorrelação (*ACF*) e da autocorrelação parcial (*PACF*), como já indicado. Sendo a correlação uma interdependência entre duas quantidades variáveis, percebe-se a autocorrelação como sendo a correlação interna das observações de uma série temporal, expressa como uma função temporal com atraso entre estas observações [14]. Ou seja, uma autocorrelação de ordem k , é uma dependência linear entre duas variáveis, sendo uma delas relativa à própria variável com um atraso k , x_{t-k} .

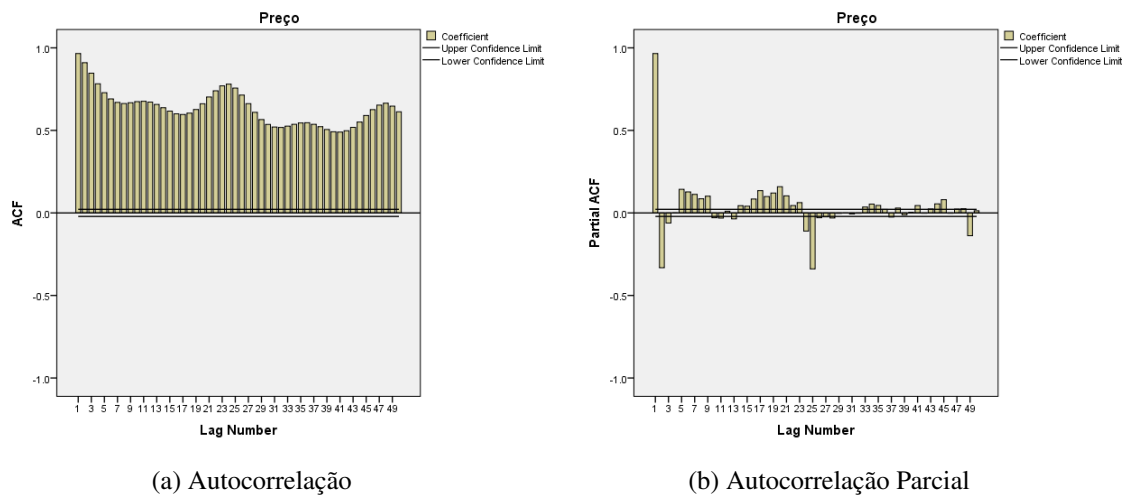


Figura 2.3: Exemplo de funções de autocorrelação e autocorrelação parcial.

Como se verifica no exemplo representado pela Figura 2.3, no caso relativo ao preço da electricidade em Portugal, nota-se uma forte correlação entre algumas variáveis (e.g. 24 e 48 horas anteriores à data). Ainda assim, se colocarmos em questão apenas a *ACF*, torna-se mais difícil discriminar modelos *AR* e daí surge a necessidade de tirar partido da *PACF* que impede que se estabeleça uma dependência em cadeia dos elementos intermédios.

Segundo [15], através de uma análise do comportamento de ambas as funções *ACF* e *PACF* consegue-se o tipo de modelo a utilizar, e determinar os parâmetros p e q . Na tabela 2.1, verificam-se os modelos a utilizar para os diferentes comportamentos de *ACF* e *PACF*.

Tabela 2.1: Tipo de modelo Box-Jenkins a utilizar para diferentes comportamentos de *ACF* e *PACF* [15] [16].

	MA(q)	AR(q)	ARMA(p, q)
ACF	Desprezável após atraso q	Decrescente	Decrescente após q
PACF	Decrescente	Desprezável após p	Decrescente após p

A abordagem proposta pelo modelo *Box-Jenkins* é então uma hipótese minimamente satisfatória para séries repetíveis ou com padrões facilmente reconhecíveis, mas peca pela falta de eficácia em cenários mais caóticos onde o atraso temporal não seja uma atração basilar no que toca à obtenção admissível de uma previsão final. Para os casos onde o reconhecimento de padrões seja apenas exequível com a adição de variáveis externas e interligadas, sugere-se a admissão de métodos de previsão mais complexos. Note-se que o modelo *Box-Jenkins* estabelece uma relação linear entre a grandeza a prever e as respetivas entradas, pelo que pode não ser tão eficaz se existirem não linearidades envolvidas na relação entre variáveis.

2.5 Redes Neurais Artificiais

As redes neurais artificiais ou, na literatura inglesa, *artificial neural networks* (ANN), são uma abordagem computacional baseada num vasto leque de unidades de processamento elementar, os *neurónios*, são modelos simplificados do sistema nervoso central do ser humano. Trata-se de uma estrutura extremamente interconectada de unidades computacionais sendo que cada ligação pode, ou não, ser inibitória do seu efeito de conexão da rede neuronal. Pretende então replicar a computação do cérebro humano, que se baseia num tipo de processamento bastante diferente daquela presente num computador convencional. Num computador, a informação é essencialmente processada em série (i.e., sequência de instruções). No cérebro, o processamento é massivamente paralelo. Assim, este tipo de redes pretende realizar operações que repliquem aquelas realizadas pelo cérebro humano. Ora, é uma unidade de processamento de informação muito completa capaz de estruturar e organizar os seus componentes estruturais, os neurónios, tornando-se capaz de executar tarefas de forma mais eficaz do que outros modelos equivalentes.

2.5.1 História

A era moderna das redes neurais começa com o trabalho pioneiro de *Warren McCulloch* e *Walter Pitts*¹ em 1943 [18].

Originalmente, o seu trabalho indica que é possível associar o conhecimento e lógica matemáticos aos estudos da neuro-ciência e conseguiram mostrar, ainda que com um modelo simplista que, com uma quantidade de unidades elementares e de conexões, constituída uma rede, poderia solucionar computacionalmente diversos tipos de problemas. A solução baseava-se num modelo equivalente a um função de transferência, capaz de simular expressões lógicas finitas, apresentando-se apenas como um dispositivo com *inputs* e *outputs* binários. Nesta solução, a cada neurónio (com diversas entradas), é imposto um determinado limite (*threshold*, θ), e recebe sinapses excitatórias com pesos equivalentes, sendo que, quando a soma das suas entradas fosse superior a esse limite, e se não recebesse nenhum sinal inibitório, era ativado o neurónio seguinte (se inativo) [18].

O passo seguinte no desenvolvimento das redes neurais, já no final dos anos 40, passa pela publicação do livro *The Organization of Behavior* por *Donald Hebb*, no qual indica que a plasticidade neuronal é capaz de garantir a aprendizagem não supervisionada pela contínua mudança na conectividade do cérebro perante a mudança e a execução de diferentes tarefas [17].

Sensivelmente 15 anos após a publicação original de McCulloch e Pitts, surge uma nova abordagem no reconhecimento de padrões, introduzida por *Rosenblatt* em 1958 [19], com o seu trabalho relativo ao *perceptron* nos cenários da aprendizagem supervisionada, demonstrando que as redes neurais poderiam ser úteis. Porém, a publicação de livros da parte de *Minsky* e *Papert*

¹McCulloch era um psiquiatra e neuroanatomista e Pitts um prodígio matemático que se juntaram para investigação em 1942 [17].

[20][21], veio demonstrar que surgem limites naquilo que um *perceptron* de uma só camada poderia calcular. Não sendo ainda possível, naquele momento, implementar modelos de múltiplas camadas, impõe-se um período de estagnação científica na área das redes neuronais, também motivada pela falta de poder computacional para acompanhar os avanços teóricos existentes.

Por volta do final dos anos 70 e no decorrer dos anos 80, surgem diversas contribuições para a teoria e design das redes neuronais, permitindo um ressurgimento do interesse no seu estudo pelo aparecimento da *teoria da ressonância adaptativa (ART)* [17]. Para além disso, com o aparecimento do treino e aprendizagem não supervisionados, alinhados com um enorme aumento do poder de processamento, as redes neuronais passaram também a ser aplicadas ao tratamento da imagem e do reconhecimento visual em larga escala.

Nos anos 80 e 90, também graças às publicações de *Hopfield* [22], e pelo desenvolvimento do algoritmo de retro-propagação pela parte de *Rumelhart*, *Hinton* e *Williams* em 1986 [23], garante-se um ressurgimento da popularidade das redes neuronais pelo interesse no treino de *MLP (multi layer perceptron)*. Este avanço foi de encontro com a informação libertada por *Werbos* na sua tese de doutoramento pela Universidade de Harvard, datada de 1974 [24].

Com isto, verificou-se uma nova "explosão" das redes neuronais e da sua aplicação em diferentes campos, o que garantiu o desenvolvimento destas e o aparecimento de novos modelos.

2.5.2 Funcionamento de uma Rede Neuronal

Numa rede neuronal, cada neurónio recebe um sinal de entrada representativo duma combinação da informação dos anteriores, e até por estímulo externo. Este sinal é processado localmente por uma ativação ou função de transferência e produz um sinal de saída já transformado para os restantes neurónios ou para uma saída final.

Ainda que individualmente cada neurónio possa ter uma implementação imperfeita da sua função, coletivamente uma rede pode desempenhar um elevado número de tarefas de forma muito eficiente. Compreende-se então que, para o processamento de informação com estas características, as *ANN* surgem como um modelo computacional muito eficaz e capaz de aprender a partir de exemplos e generalizar outros nunca antes executados [25]. São sistemas capazes de realizar aprendizagem autónoma bem como reconhecimento de padrões, treináveis e não diretamente programáveis, sobressaindo-se em aplicações onde a solução dificilmente é alcançada em programas e modelos normais. Visto de um prisma de uma máquina adaptativa, uma *ANN* pode ser definida como [17]:

Uma *ANN* é um processador de funcionamento distribuído e paralelo, composto por unidades de processamento elementares, com a propensão natural para armazenar conhecimento empírico e torna-lo acessível ao utilizador, aproximando-se do comportamento do cérebro em dois aspetos:

- O conhecimento é adquirido perante um dado ambiente, passado por um processo de aprendizagem.

- Este conhecimento é armazenado em conexões, ligações ou sinapses, entre neurónios.

Durante este procedimento, existe um algoritmo de treino monitorizado pelo peso das conexões que é ajustado de forma a que o objetivo final seja atingido. Ainda que esta seja a habitual composição de uma *ANN*, a sua topologia também pode ser alterada, indo de encontro às possíveis novas conexões existentes no cérebro humano. Portanto, temos um sistema bem sustentado e capaz, caracterizado por uma estrutura imensamente complexa, não linear e paralela, capaz de organizar os seus constituintes.

Uma *ANN* apresenta um poder computacional que se sustenta sobre dois grandes pilares: realça-se uma topologia característica que pretende salientar o paralelismo e também a sua capacidade de generalizar e aprender. Ora, pretende-se com isto garantir uma resposta adequada a novos cenários mas claro, ter a informação e as experiências passadas como parâmetro basilar na sua execução. Estas duas características são particularmente interessantes uma vez que garantem que as redes neuronais triunfem em situação que de outra forma seriam inexequíveis. Note-se que por si só, uma rede neuronal poderá, evidentemente, não ser capaz de solucionar um problema muito complexo. Daí, estas redes são mutuamente alicerçadas a outros sistemas capazes, permitindo uma execução mais completa e eficaz. Mesmo afixada uma enorme complexidade numa *ANN*, é certo afirmar que a sofisticação de processos do cérebro humano ainda está bastante longe de ser alcançada, pelo menos num futuro próximo.

Ainda assim, uma *ANN*, com uma estrutura massivamente distribuída e paralela, é capaz de generalizar. Quando falamos de generalização, falamos da sua capacidade de extrair resultados razoáveis (*outputs*) para entradas (*inputs*) desconhecidas, isto é, não visualizadas, no processo de treino e aprendizagem. Isto é, a rede deve ser capaz de permitir a estimação de flutuações mesmo quando estes não ocorreram e não foram inseridos como informação no processo de aprendizagem. Contudo, mais uma vez, importa referir que, mesmo com estas valências, uma *ANN* ainda está longe de ser perfeita, e deve-se salientar a sua necessidade de trabalho paralelo com outras ferramentas.

Na sua globalidade, uma *ANN* é eficaz, mas devem-se destacar alguns benefícios na sua utilização, características únicas dificilmente encontradas em outros mecanismos ou técnicas [17][26].

1. **Não-Linearidade;** i.e., um neurónio pode ser linear ou não linear. Porém, uma rede neuronal constituída pela interconexão de neurónios não lineares, é por si só não linear. Ainda mais, trata-se de uma espécie distinta de não-linearidade, particularmente por ser distribuída ao longo de toda a rede. A não-linearidade é uma propriedade de capital importância se o mecanismo responsável pela criação do sinal de entrada (e.g. sinal sonoro), for intrinsecamente não linear.
2. **Mapeamento Input-Output;** i.e., apresentado um conjunto de informação de entrada, e um respetivo conjunto de saída, diretamente relacionáveis, a rede assume-os como uma entrada única e uma correspondente saída esperada. Ora, sendo a rede estabelecida com um dado exemplo aleatório desse conjunto, os pesos das conexões, ou das ligações sinápticas

são ajustados de forma a que a diferença entre o sinal produzido e o real, correspondente à entrada de exemplo, seja a menor possível, e de acordo com os critérios estatísticos apropriados.

3. **Adaptabilidade;** i.e., as redes neuronais têm incorporada a capacidade de adaptar os pesos das conexões a mudança no ambiente envolvente. Particularmente, uma rede neuronal treinada para operar num dado cenário específico, pode facilmente ser re-treinada para conseguir lidar com pequenas mudanças nas condições de operação. Ainda mais, quando opera sob um ambiente não estacionário (i.e. onde as estatísticas são vulneráveis à mudança no tempo), uma rede neuronal pode ser configurada para alterar os pesos à medida da mudança, em tempo real. A arquitetura natural de uma rede neuronal para reconhecimento de padrões, processamento de sinais e controlo de aplicações, acoplado com esta capacidade de adaptabilidade, torna-a numa ferramenta útil em processos adaptativos de classificação de padrões, de processamento de sinal e de controlo. Como regra geral, é possível afirmar que, quanto mais adaptativo tornamos um certo sistema, garantindo em todas as ocasiões que este se mantém estável, mais robusta se torna a sua performance quando este sistema for colocado a operar num ambiente não estacionário. Note-se que, a adaptabilidade por si só, não leva sempre à robustez, mas por outro lado poderá causar o contrário. Vejamos que o caso de um sistema adaptativo, com constantes de tempo baixas, poderá rapidamente sofrer alterações e tenderá a responder a flutuações ocasionais causando uma degradação drástica no desempenho do sistema. De forma a compreender completamente os benefícios da adaptabilidade, as constantes de tempo de um sistema deverão ser suficientemente longas para que este ignore as flutuações espúrias e ainda assim ser capaz de responder a mudanças no ambiente em que opera.
4. **Resposta evidencial;** i.e. visualizada num cenário de classificação de padrões, uma rede neuronal pode ser projetada para fornecer informação não só apenas acerca de um padrão em particular, mas também indicar a confiança na decisão tomada. Isto permite que tenhamos a possibilidade de rejeitar respostas ambíguas tornando o desempenho da classificação mais eficaz.
5. **Informação contextualizada;** i.e. o conhecimento é representado pela estrutura e pelo estado de ativação da rede neuronal. Cada neurónio tem o potencial de ser afetado pela atividade global dos restantes e consequentemente a informação para um certo contexto é naturalmente lidada pela rede neuronal.
6. **Tolerância ao erro;** i.e. robustez e degradação suave. Uma rede neuronal tem o potencial para tolerar erros nos dados e é capaz de realizar processos computacionais robustos, tendo em conta que a sua performance se degrada se presente perante condições de operação adversas. Ou seja, uma rede neuronal permite processar ruído ou informação incompleta de forma eficiente, e mantém a sua performance mesmo com a desativação de algumas das conexões entre neurónios.

7. **Analogia neurobiológica;** i.e. A estrutura de uma rede neuronal é motivada pela analogia com o cérebro humano, o que é uma prova de que a tolerância ao erro no processamento paralelo é não só fisicamente possível, mas também rápido e poderoso. Os neurobiólogos tendem a considerar as redes neuronais como uma ferramenta de investigação para a interpretação de fenómenos neurobiológicos. Por outro lado, os engenheiros olham para a neurobiologia como fonte de novas ideias para a resolução de problemas mais complexos do que os convencionais que são resolvidos por outras técnicas convencionais.

2.5.3 Arquitetura de uma Rede Neuronal

O elemento mais elementar da constituição de uma rede neuronal é o neurónio, ou nodo. O termo nodo, ou simplesmente unidade, é habitualmente usado para que se consiga distinguir entre um neurónio natural e artificial [26]. Um neurónio apresenta um funcionamento simples e como tal, comporta-se como um comparador que faz produzir uma determinada saída consoante o efeito das diversas entradas exceda um certo limite.

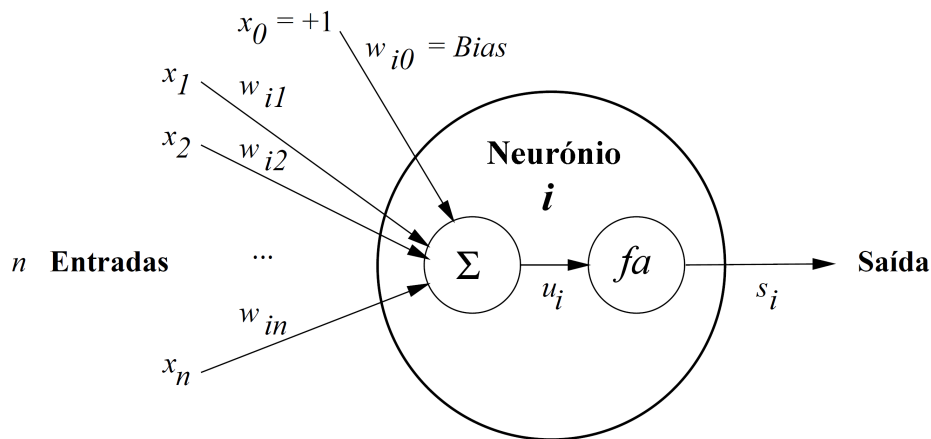


Figura 2.4: Estrutura geral de um neurónio [26].

De acordo com a Figura 2.4, um neurónio é composto por [26]:

- Um *conjunto de conexões* (w_{ij}), cada qual etiquetada por um *peso*; i.e. um número real ou binário (mais comum na forma real) que impõe um efeito tipicamente excitatório para valores positivos e inibitório para valores negativos. Ora, o sinal de entrada (x_j) é multiplicado pelo peso correspondente, w_{ij} , identificando i como o neurónio que está a ser ponderado, e j como o neurónio remetente. Ainda mais, existe a possibilidade da existência de uma conexão extra, *bias*, que estabelece uma dada tendência ou inclinação em todo o processo; i.e., adiciona uma constante (w_{i0}) de forma a que sejam impostas as condições operacionais indicadas para o neurónio em questão [27].

- Um *integrador* (g), com a função de agregar os n argumentos de entrada (estímulos) a um único valor. Habitualmente opta-se por utilizar a função de soma (Σ) de forma a pesar todas os *inputs* numa combinação linear.
- Uma *função de ativação* (f), que tem a possibilidade de condicionar o sinal de saída, embutindo um elemento de não linearidade ao processo.

Assim, pode-se reduzir o funcionamento do neurónio a:

$$u_i = g(1 \cdot w_{i0}, x_1 \cdot w_{i1}, x_2 \cdot w_{i2}, \dots, x_n \cdot w_{in}) \quad (2.3)$$

$$s_i = f(u_i) \quad (2.4)$$

para um neurónio i com n entradas e apenas uma saída, com u_i a representar o ganho do neurónio i e s_i a saída do neurónio.

De notar que a função de ativação é relevante, até pelas diferentes topologias que pode tomar (e.g. *Heaviside*, linear, quadrada, gaussiana, etc.) como pelo efeito que causam no sinal de saída. Assim, as funções de ativação mais comuns numa ANN são:

Tabela 2.2: Tipos de funções de ativação

Nome	Função f	Contradomínio
Limiar	$\begin{cases} 1, & u_i \geq 0 \\ 0, & u_i < 0 \end{cases}$	$\{0, 1\}$
Linear	u_i	$] -\infty, +\infty[$
Por troços	$\begin{cases} 1, & u_i \geq 0.5 \\ ku_i, & -0.5 < u_i < 0.5 \\ 0, & u_i \leq -0.5 \end{cases}$	$[0, 1]$
Logística	$\frac{1}{1+e^{-ku_i}}$	$[0, 1]$
Tangente hiperbólica	$\tanh(ku_i)$	$[-1, 1]$
Sin	$\sin(u_i \bmod 2\pi)$	$[-1, 1]$
Cos	$\cos(u_i \bmod 2\pi)$	$[-1, 1]$
Gaussiana	$e^{\frac{-u^2}{2k^2}}$	$[-1, 1]$
Quadrada	$-\text{sign}(u_i)u^2$	$] -\infty, +\infty[$

A tabela 2.2 indica alguns dos tipos mais comuns de funções de ativação usadas nas redes neuronais, sendo que k representa a inclinação da função, *mod* o resto de uma divisão inteira e $\text{sign}(x) = \frac{x}{|x|}$. A primeira função, do limiar ou de *Heaviside*, é usada habitualmente nos nodos de

McCulloch-Pitts [18], onde a saída toma o valor $+1$ se o ganho for não negativo, retomando a filosofia do *tudo-ou-nada* [26]. De notar as funções que apresentam não linearidade, destacando claramente a função logística, uma função *sigmoide*. Esta, com uma forma de curva idêntica a um *S*, é de forma destacada, uma das funções mais utilizadas nas *ANN*. Ora, trata-se de uma função crescente e que aposta numa divisão imperial acerca da sua componente linear e não linear (Figura 2.4) [28]. De notar também a função *tanh*, uma outra função sigmoide (em escala diferente), também muito usada em redes neuronais. Uma variação da inclinação k permite obter curvas de diferentes declives, sendo que, no seu limite, com $k \rightarrow \infty$, a função tenderá para um comportamento equivalente à função limiar.

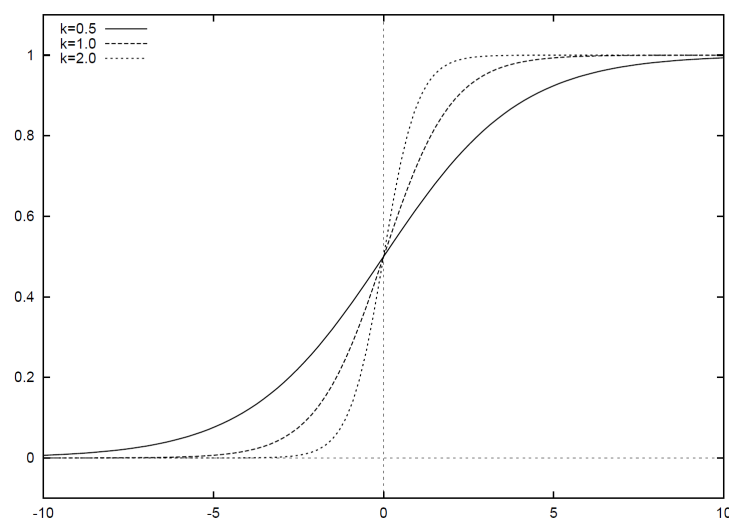


Figura 2.5: Exemplo da função sigmoide para as inclinações de $k = 0.5$, $k = 1.0$ e $k = 2.0$

Compreendida a arquitetura de um neurónio, importa perceber o tipo de conexões que estes podem estabelecer entre si. Um nodo pode estar ligado a centenas, ou mais, de outros neurónios e tendem a agrupar-se em camadas, pelo que se consegue distinguir três distintos tipos de ligação:

- *convergentes*; i.e., vários neurónios estão conectados a um único neurónio;
- *divergentes*; i.e., um só neurónio pode estar ligado a diversos neurónios através de uma arborização radial;
- *encadeadas ou cíclicas*; i.e., ligações que podem envolver vários neurónios formando encaamentos ou ciclos.

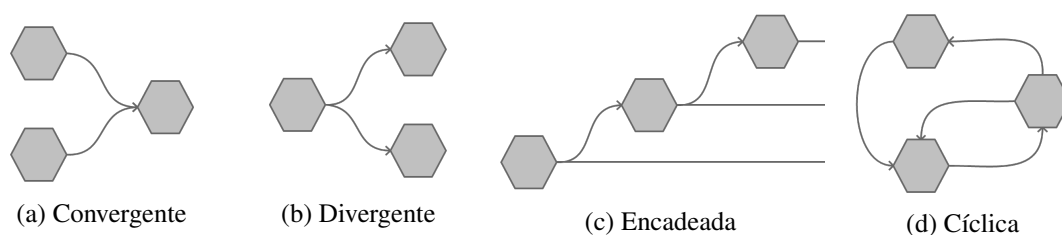


Figura 2.6: Diferenças entre os modos de ligação entre neurónios.

A forma como os neurónios estão dispostos, ou a forma como a rede neuronal é estruturada está intimamente relacionada com o algoritmo de aprendizagem escolhido para treinar a rede. Assim, surge o conceito de aprendizagem ou de algoritmos de aprendizagem, que são usados para arquitetar uma rede neuronal. Ainda que não tenham sido mencionados os algoritmos de aprendizagem, é relevante considerar as arquiteturas ou estruturas das redes como ponto de partida. Em regra geral, são identificados três tipos fundamentais de classes de redes neuronais [17].

1. **Redes *Feed-Forward* de uma camada:** Numa rede neuronal com camadas, os neurónios estão organizados sob forma de camadas. Na forma mais singela de uma rede com camadas, existe uma camada de inputs que se projetam sobre uma outra camada de outputs, mas apenas neste sentido. Por outras palavras, é uma rede estritamente *Feed-Forward*, ou seja, é acíclica. Como é possível verificar pela Figura 2.7, estamos presentes a uma rede de uma única camada, onde a camada de entrada não é considerada, já que nesta não é feito qualquer processamento computacional.

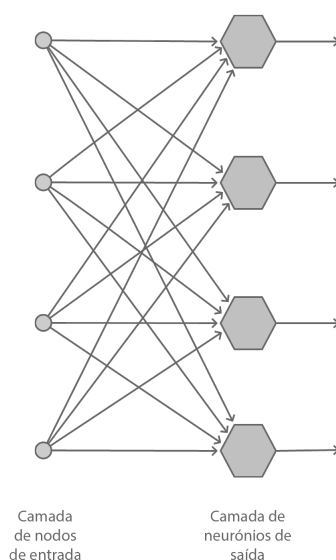


Figura 2.7: Rede neuronal *Feed-Forward* de apenas uma camada.

2. **Redes *Feed-Forward* de múltiplas camadas:** A segunda classe de redes *Feed-Forward* distingue-se pela presença de uma ou mais camadas escondidas, cujos membros correspondentes são denominados por unidades escondidas. A função destas unidades, destes neurónios, é a de intervir entre a entrada externa e a saída da rede de uma forma útil. Ao adicionar uma ou mais camadas escondidas, a rede passa a ser capaz de extrair relações de complexidade de maior ordem. De certa forma, a rede passa a ter uma perspetiva global mesmo tendo em conta as suas conexões locais, devido ao aumento de ligações sinápticas e pelo aumento da dimensão das interações neuronais. A habilidade das unidades escondidas em extrair informação de ordem de grandeza superior é particularmente valiosa quando a dimensão da camada de entrada é substancial. Os nodos da camada de entrada fornecem elementos pertencentes ao padrão de ativação, constituindo respetivamente, por sua vez, também eles um conjunto de entradas para os neurónios da segunda camada (i.e., da primeira camada escondida). Os sinais de saída desta segunda camada vão passar a ser os sinais de entrada da terceira, e assim sucessivamente pelo resto da rede. Tipicamente, os neurónios de cada camada têm como entradas, os sinais de saída apenas da camada precedente. O conjunto de sinais de saída da última camada da rede constitui, finalmente, a saída global da rede. A Figura 2.8 ilustra o esquema duma rede neuronal *Feed-Forward* de múltiplas camadas, para o caso com uma camada escondida e uma camada de saída. Refere-se ao caso de uma rede 10-4-2 uma vez que contém 10 nodos de entrada, 4 unidades (ou neurónios) escondidos e 2 neurónios de saída. Ora, uma rede com m nodos de entrada, h_1 neurónios na primeira camada escondida, h_2 neurónios na segunda camada escondida e q neurónios na camada de saída, é uma rede $m - h_1 - h_2 - q$.

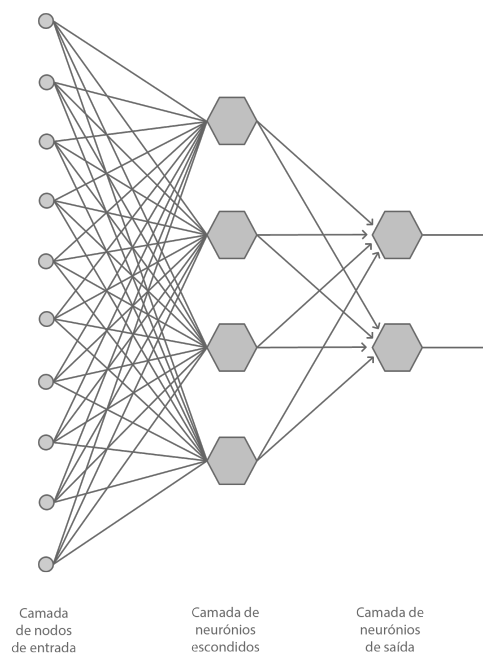


Figura 2.8: Rede neuronal *Feed-Forward* de múltiplas camadas.

A rede neuronal identificada pela Figura 2.8 é considerada como completamente conectada uma vez que cada neurónio de cada camada está conectado a todos os neurónios subsequentes das camadas adjacentes. Nos casos em que umas das ligações mencionadas não existe, a rede diz-se parcialmente conectada.

3. **Redes recorrentes:** Uma rede recorrente, distingue-se das anteriores pela presença de pelo menos um ciclo com *feedback*. Por exemplo, uma rede recorrente pode consistir numa única camada de neurónios realimentados entre si.

2.5.4 Aprendizagem de uma Rede Neuronal

Um dos principais atributos de uma *ANN* é a sua plasticidade no processo de treino, a sua capacidade de reagir à mudança e de aprender segundo as alterações que o seu ambiente proporciona. Uma rede neuronal é então capaz de aprender, de melhorar o seu desempenho, através de um processo interno de ajustamento aplicado ao peso das sinapses e dos níveis da *bias*. Ainda que não seja trivial definir o conceito de aprendizagem de uma rede neuronal, justificado pela presença de diferentes pontos de vista, segundo [17], a aprendizagem de uma *ANN* envolve uma sequência de etapas:

1. A rede neuronal é *estimulada* por um certo ambiente.
2. A rede neuronal proporciona alterações nos parâmetros livres (i.e. pesos das ligações) como resultado desta estimulação.
3. A rede neuronal responde de uma nova forma ao ambiente devido às mudanças que ocorreram na sua estrutura ou arquitetura interna.

Ora, uma *ANN* realiza o seu processo de aprendizagem tendo como base um bom conjunto de regras, de procedimentos e paradigmas matemáticos a respeitar, introduzindo o conceito de *algoritmo de treino*. Cada algoritmo apresenta fundamentos distintos, características específicas e performances diferentes, surgindo a necessidade de optar entre estes, dependendo do cenário a trabalhar. Cada algoritmo apresenta as suas vantagens e desvantagens, distinguidos pelo seu paradigma ou pelo ambiente e sob as circunstâncias em que a rede deve operar. Uma *ANN*, de uma forma geral, pode funcionar sob três paradigmas de aprendizagem [29]:

1. **Supervisionada:** Este paradigma baseia-se na existência de um "professor", isto é, já são fornecidos valores corretos à rede. Ora, a rede vai realizar o processo de aprendizagem perante um conjunto de padrões (P), onde cada padrão (p), também denominado *exemplo* ou *conjunto de treino*, é composto por um vetor de entrada (x^p) e por um vetor de saída (s^p). Neste processo, realiza-se uma comparação entre o valor desejado (t^p) e o valor de saída da rede, onde se calcula um erro ($e^p = t^p \ominus y^p$, com \ominus a ser a função de erro). Assim, o erro calculado é usado para reajustar os pesos das ligações da rede, com objetivo de o reduzir.

Uma iteração do algoritmo de treino é formada por um conjunto de etapas, de ajustes iterativos para todos os casos do conjunto de treino. Desta forma, e dado o conjunto de valores em questão, o processo de aprendizagem conclui-se quando, após algumas iterações, o valor do erro é reduzido para valores considerados aceitáveis.

2. **De Reforço:** Ainda que neste segundo paradigma também se considere a presença de um "professor", a os valores corretos não são inseridos na rede. Por outro lado, é fornecida uma indicação apenas sobre se a saída da rede toma valores corretos, ou não, usando esta informação para melhorar a sua performance. Desta forma, caso a resposta seja acertada, é reforçado o peso dessas ligações e reduzido para a situação inversa.
3. **Não Supervisionada:** Neste caso, estamos perante uma abordagem totalmente distinta, onde não são impostas indicações acerca do desempenho da saída. Assim, o processo de aprendizagem compreende um trajeto onde é realizada uma descoberta das características dos valores de entrada, agregando padrões e características estatísticas do conjunto de treino. Neste caso, a ideia essencial é agrupar os exemplos de entrada em blocos (*clusters*), ou seja, dividir estes padrões em conjuntos cujos elementos têm características similares.

Para além dos paradigmas de aprendizagem, é possível dissolver este processo em cinco diferentes modelos. Ora, segundo [17], a aprendizagem deve responder a uma das cinco regras de aprendizagem:

1. *Hebbian*. Para além de ser o modelo mais antigo, é um dos modelos mais utilizados fora da engenharia (onde reina provavelmente o gradiente descendente e algoritmos derivados). Foi proposto por *Hebb*²

“Quando um axónio da célula *A* está demasiado próximo para excitar uma célula *B* e dispara-a de forma repetitiva ou persistente, então ocorre um qualquer processo metabólico numa ou ambas células de forma a aumentar a eficiência da célula *A* ao disparar *B*” [30].

Esta afirmação, envolvido num contexto neuro biológico, foi expandido para uma regra de duas partes:

- Se dois neurónios, de ambas as extremidades de uma sinapse são ativados simultaneamente (i.e., de forma síncrona), então, a força dessa sinapse é seletivamente aumentada.
- Se dois neurónios, de ambas as extremidades de uma sinapse não são ativados de forma síncrona, então a sinapse é seletivamente enfraquecida ou eliminada.

Este tipo de conexão, então designada de sinapse *Hebbian*, depende do tempo e localização espacial e é muito utilizada na aprendizagem não supervisionada.

²Donald Hebb (1904-1985): Psicólogo canadiano. Teórico influente na relação do cérebro com o comportamento.

2. *Competitiva*. Neste modelo, as saídas dos neurónios da mesma camada realizam uma competição entre si de forma a que se tornem mais ativas, resultando num neurónio a ser ativado num dado instante, com base na estratégia do *vencedor-fica-com-tudo*. No início do processo, os neurónios de uma certa camada, possuem conexões com pesos pequenos e desiguais. Quando é imposto um padrão à rede, um dos elementos dessa camada responderá melhor do que os restantes, sendo “premiado” com um reforço no peso das suas ligações. Esta regra é muito apropriada para processos como a *extração de características*, na resolução de problemas de classificação.
3. *Estocástica*. Este tipo de aprendizagem caracteriza-se pelo ajuste dos pesos de uma forma probabilística. Temos como exemplo o *Simulated Annealing* aplicado a máquinas de *Chau-chy* e *Boltzmann* com o estado dos neurónios a serem determinados por uma distribuição probabilística [27].
4. *Baseada na Memória*. Esta aprendizagem tem como base o armazenamento de todas, ou quase todas, as experiências passadas, que são armazenadas numa memória de pares entrada-saída. No surgimento de um novo vetor x^p , nunca antes introduzido na rede, o algoritmo de aprendizagem responde pela procura de vetores na região vizinha de x^p . Por norma, este algoritmo é composto por duas componentes:
 - um critério de definição da vizinhança local do vetor x^p ;
 - uma regra de aprendizagem aplicada ao conjunto de treino da vizinhança local de x^p .

Este tipo de aprendizagem é utilizado, por exemplo, nas redes *Radial-Basis Functions (RBF)* [31].

5. *Gradiente Descendente*. Como referido, a aprendizagem supervisionada pretende reduzir o erro entre o valor pretendido e o valor de saída da rede neuronal. Perante isto, neste modelo, o erro assume um papel de mecanismo de controlo, garantindo ajustamentos e daí, melhores respostas. Tem como principal objetivo minimizar uma dada função de custo, ξ , definida em termos do sinal de erro e^p . Tem uma regra de aprendizagem, conhecida como regra *delta* ou *Widrow-Hoff*, baseada na resolução da seguinte equação [32]:

$$\Delta\omega = \eta \nabla \xi \quad (2.5)$$

onde η representa uma constante positiva denominada *taxa de aprendizagem*, e $\nabla \xi$ o gradiente da função de custo. É uma aprendizagem muito utilizada e reconhecida, muitas vezes associada a redes de múltiplas camadas e ao algoritmo de *Back-Propagation* [33].

2.5.5 Algoritmo de Treino *Levenberg-Marquardt*

Um dos algoritmos de treino de redes neuronais mais utilizados tem como base o algoritmo *Levenberg-Marquardt (LMA)*. É um método caracterizado pela retropropagação e habitualmente

uma das primeiras opções de escolha dos métodos supervisionados pela sua rapidez de treino de redes *Feed-Forward*, ainda que requeira um maior uso de memória, quando comparado com outros algoritmos.

Inicialmente publicado em 1944 por *Kenneth Levenberg* [34], e re-descoberto em 1963 por *Donald Marquardt* [35], o *LMA* é um método normalmente utilizado para solucionar problemas não lineares de mínimos quadrados. O *LMA* é assim usado para resolver problemas genéricos de ajuste de curvas ainda que, tal como muitos outros algoritmos, este decifra apenas um mínimo local, que não é necessariamente o mínimo global.

O *LMA* é então uma interpolação entre o método de *Gauss-Newton* (*GNA*) e o método do gradiente, ainda que mais robusto que o primeiro. Percebe-se isto quando, na maioria dos casos, é possível obter uma solução mesmo quando se parte de um ponto muito longe do mínimo ideal, ainda que mais lento que o método de *Gauss-Newton*. Dadas as semelhanças, o *LMA* pode ser visto como uma aproximação do *GNA* mas que é aplicado apenas num intervalo de confiança.

Segundo [36], existem variadas técnicas de otimização numéricas que podem ser usadas para acelerar o processo para obtenção de convergência em algoritmos de aprendizagem por retro-propagação, sendo o *LMA*, um dos métodos que é passível de ser integrado nestes modelos. Testado em diversos problemas com funções de aproximação, e comparado com outros métodos, os resultados indicam que o *LMA* é muito eficiente no treino de *ANN* compostas por poucas centenas de pesos. Ainda que eleve o poder de processamento requerido para cada iteração do método, a eficiência resultante é capaz de se sobressair, principalmente quando é requerido um elevado nível de precisão. Para além disso, o *LMA* é capaz de convergir para situações onde outros métodos falham, garantindo ser uma solução capaz de envergar problemas mais complexos.

De forma semelhante aos métodos quase-Newton, o *LMA* foi desenvolvido para ter uma abordagem próxima ao treino de segunda ordem sem ter de computar a matriz *Hessiana*. Ora, a função de performance toma a forma da soma de quadrados (como é tipicamente realizado no treino de redes *Feed-Forward*), com a matriz *Hessiana* a ser aproximada como [37]:

$$H = J^T J \quad (2.6)$$

com o gradiente a ser calculado como:

$$g = J^T e \quad (2.7)$$

onde *J* representa a matriz *Jacobiana* que contém as primeiras derivadas dos erros da rede considerando os pesos e as *bias* e onde *e* representa o vetor dos erros da rede. De notar que a matriz *Jacobiana* pode ser calculada através de métodos de retro-propagação clássica [36], processo bastante menos complexo do que o cálculo completo da matriz *Hessiana*. O *LMA* utiliza então uma aproximação da matriz *Hessiana* por um processo idêntico ao da atualização de *Newton*:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (2.8)$$

Assim, quando o escalar μ toma o valor de zero, estamos perante o método de *Newton*, usando a aproximação da matriz *Hessiana*. Quando μ toma valores elevados, estamos em face do método do gradiente com um degrau reduzido. Esclareça-se que o método de *Newton* é mais rápido e preciso quanto perto de um erro mínimo, pelo que se deve transitar para este o mais rápido quanto possível. Ora, μ deve ser reduzido, após cada passo com sucesso (i.e. aplica-se uma redução na função de desempenho), ou aumentado apenas quando uma tentativa aumenta a função de desempenho. Resumindo, percebe-se que, neste caso, a função de desempenho é sempre reduzida a cada iteração do algoritmo.

2.5.6 Aplicações de Redes Neurais

Esclarecidos os temas relativos à arquitetura da rede e métodos de aprendizagem, importa compreender que estes dois parâmetros são claramente influenciados pela tarefa ou função a desempenhar pela ANN. Ora, podem-se estratificar as possíveis tarefas de uma rede da seguinte forma [29][17]:

1. *Memória Associativa*. É um tipo de memória distribuída com aprendizagem por associação. Pode tomar duas formas: *autoassociação* ou *heteroassociação*. Acerca do primeiro, a rede armazena um conjunto de padrões sendo posteriormente apresentado uma versão distorcida do padrão original, ao qual deve ser devolvido o padrão original. No segundo caso, os padrões estão organizados por pares de entrada e saída. Neste caso, estamos perante um caso de aprendizagem supervisionado, contrariamente ao desenvolvido no caso autoassociativo, onde a aprendizagem é não supervisionada.
2. *Diagnóstico*. É uma tarefa muito comum em diferentes cenários como a engenharia, medicina ou até na produção. Trata-se na verdade de um processo de classificação; i.e., requer uma correta associação entre entradas, que representem dados indicadores de um estado (e.g. sintomas ou comportamento anormal), com o correspondente diagnóstico (e.g. doença ou falha de equipamento). Habitualmente, as ANN são integradas em sistemas periciais complexos, envolvendo também outros paradigmas (e.g. *aprendizagem por regras*).
3. *Reconhecimento de padrões*. Formalmente, este processo é definido como sendo aquele através do qual um sinal ou padrão recebido é associado a uma de diversas categorias possíveis. Ora, inicialmente, é imperial treinar uma rede onde os padrões, associados a uma mesma categoria, são alimentados à rede, repetitivamente. Mais tarde, fornece-se um novo padrão à rede que deverá ser capaz de assinalar a categoria correta, com base na informação previamente assegurada. O reconhecimento de padrões por uma rede neuronal pode tomar duas formas. A forma mais simples implica uma rede *Feed-Forward* de múltiplas camadas, com treino supervisionado, com os neurónios intermédios a assegurar uma extração de características; i.e., uma transformação da entrada (x) num ponto intermédio (y) pertencente a uma dimensão inferior. Com esta redução, o processo de classificação torna-se mais simples e descrita como sendo uma transformação do ponto intermédio (y) numa das classes

possíveis num espaço de decisão r -dimensional, para r classes distintas. A segunda forma, uma máquina de aprendizagem, é constituída por duas secções, a primeira para extração de características por treino não supervisionado e o restante para classificação, por treino supervisionado. De notar ainda que as *ANN* são redes muito eficazes na aprendizagem de processos percetivos como o processamento de imagem e reconhecimento de voz.

4. *Regressão/Previsão*. O objetivo é executar uma *ANN* capaz de modelar uma função desconhecida $f(\cdot)$ que se aproxime de $F(\cdot)$ dada por um conjunto de vetores etiquetados; i.e., um conjunto de informação composto por um par entrada saída ($x \rightarrow y$), garantido que a sua distância *euclidiana* seja reduzida o suficiente para todas as entradas, i.e.:

$$\forall x, \|F(x) - f(x)\| < \rho \quad (2.9)$$

onde ρ representa um dado valor reduzido. A regressão por si só é um procedimento razoável para a aprendizagem supervisionada. A previsão, um caso particular da regressão, é um conceito usado, entre outros, na área da economia, vendas ou produção onde se pretende estimar, até “adivinhar” os valores a surgir no futuro, com base nos valores do passado. As *ANN* têm desempenhado um papel imperial e muito eficaz como uma ferramenta de previsão, realçando-se a ocorrência (ou não) de eventos e/ou a intensidade destes.

5. *Controlo*. Esta tarefa rodeia-se de procedimentos considerados críticos e fundamentais de um certo sistema e que devem ser mantidos numa condição controlada. Assim, um *controlador* deve ser capaz de fornecer um conjunto de informação, de sinais adequados para um certo sistema, garantindo que a saída (y) acompanhe uma determinada entrada de referência (x). De referir os exemplos de veículos autónomos, robôs ou até processos de fabrico, onde as *ANN* têm sido utilizadas frequentemente e com enorme sucesso.
6. *Otimização*. Como é sabido, certos problemas que devem garantir a procura de soluções ótimas e que têm em consideração diversos parâmetros mas com um intervalo de procura demasiado extenso, até onde a exaustão de todas as alternativas se torna incomportável em termos computacionais. Uma solução para este tipo de problemas passa pelo recurso a heurísticas ou máquinas de aprendizagem (*ANN* incluídas) que devem reduzir o intervalo de pesquisa impondo possivelmente soluções mais aceitáveis.
7. *Filtragem e Compressão de dados*. De uma forma habitual, o termo filtragem remete para um processo, algoritmo ou dispositivo que extrai informação de interesse a partir de um conjunto “bruto” de dados que deve conter ruído. Por outro lado, a compressão envolve uma redução de um dado espaço n -dimensional para um outro m -dimensional, com $n > m$. São ambos evidentemente importantes para cenários repletos de dados, com conjuntos enormes de informação que deve ser processada.

2.6 O Setor Elétrico em Portugal

2.6.1 Preço da Eletricidade

O setor elétrico tem tomado um papel cada vez mais preponderante na vida de todos os cidadãos. Facilmente se entende o futuro será provavelmente mais elétrico, até porque a crescente possibilidade de uma gestão eficaz da energia elétrica garante que se coloquem de parte, cada vez mais, as energias mais poluentes. Ora, de uma forma geral, uma gestão do preço da eletricidade impõe-se como uma ferramenta de capital importância para a sociedade, no caso a indústria e até o utilizador doméstico, querendo garantir um papel basilar numa competitividade aberta e globalizada.

Nesse contexto, é executada a ideia da construção do Mercado Interior da Eletricidade da União Europeia, que arrancou a meados dos anos 90, com o objetivo de transitar o setor elétrico para uma liberalização de mercado, garantindo um melhor preço e serviço. Desta forma, foram sendo criados mercados organizados nas várias regiões da Europa. No caso, o mercado *spot* da península ibérica é gerido pela OMIE (Operador de Mercado Ibérico de Energia - Polo Espanhol), querendo transformar este mercado numa base idêntica ao que já acontece nos países nórdicos, com o *Nord Pool Spot*, entre outros.

O Mercado Ibérico de Eletricidade (MIBEL) apresenta-se como um dos mais líquidos da Europa, com preços comparáveis aos dos restantes mercados. Na verdade, de uma forma geral, o mercado ibérico apresentou até preços médios inferiores aos da média europeia (i.e., dos mercados mais relevantes). Ainda mais, importa referir que os preços deste mercado ibérico podem oscilar, salvo exceção, entre os 0 e os 180 €/MWh, enquanto que o preço europeu se movimenta num espectro de preços amplo, até aos 3000 €/MWh [38].

Neste mercado ibérico, gerido pela OMIE, e também no resto da Europa, os preços da eletricidade são estipulados diariamente pelas 12h00, para as vinte e quatro horas do dia seguinte, naquilo que é chamado por mercado diário. Garante-se que o preço e o volume de energia numa dada hora são estabelecidos pelo cruzamento entre a oferta e a procura, respeitando o modelo adotado pela UE e com base num algoritmo aprovado por todos os mercados (*EUPHEMIA*).

Os compradores e vendedores, independentemente de estarem em Espanha ou em Portugal, podem utilizar o mercado. As ofertas de compra e venda são aceites tendo em conta o peso da sua proposta, até que a interligação entre Espanha e Portugal fique totalmente ocupada. Se, numa determinada hora do dia, a capacidade da interligação for suficiente para permitir o fluxo de eletricidade negociado pelos agentes, o preço da eletricidade nessa hora será o mesmo em Espanha e em Portugal. Se, pelo contrário, nessa hora a interligação estiver lotada, nesse momento, o algoritmo para a fixação do preço é executado de forma independente, aplicando uma diferença de preços entre ambos os países. De notar mesmo assim que, durante o ano 2014, em 90% do tempo, o preço da eletricidade foi o mesmo em Portugal e Espanha, confirmando o funcionamento eficaz

da integração do mercado ibérico [38].

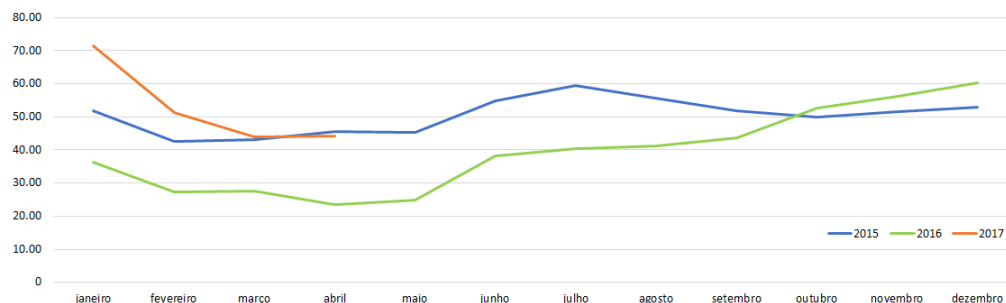


Figura 2.9: Preço médio anual do mercado diário desde 2015 (€/MWh).

Através da informação disponível em [39], e verificando a Figura 2.9, consegue-se perceber que, desde 2015, o preço médio da eletricidade estabelece-se numa gama entre os 20 €/MWh e os 75 €/MWh. Mais precisamente, uma média de 50.37 €/MWh em 2015, 39.38 €/MWh em 2016 e 52.77 €/MWh em 2017 (até abril, inclusive). De uma forma geral, desde 2015, o preço da eletricidade apresenta-se com uma média de 47.50 €/MWh.

Ainda que de uma forma anual, com valores mensais, a evolução do preço da eletricidade pareça respeitar uma certa tendência natural (i.e., uma descida no início do ano, e uma subida por volta de maio a junho), vista ao pormenor, com valores horários, percebe-se o o preço da eletricidade implica dependências extraordinárias.

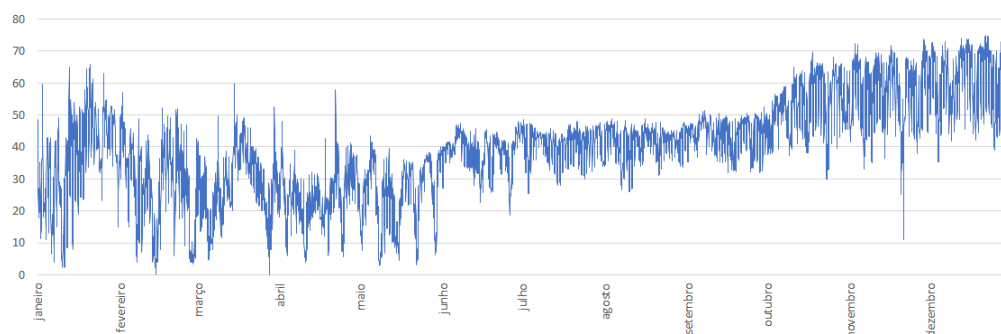


Figura 2.10: Preço médio do mercado diário de 2016 (€/MWh).

Então, como se consegue entender pela Figura 2.10, criada com dados de [39], visto de um prisma diário, constata-se que o preço apresenta alguma flutuação, que deve ser analisada.

2.6.2 Consumo de Energia Elétrica

A produção de eletricidade tem como grande pilar o contexto da concorrência e pode ser dividida sobre dois regimes legais: Produção em Regime Ordinário (PRO) e Produção em Regime Especial (PRE). A primeira relata a produção de eletricidade com base em fontes convencionais não renováveis com a PRE a representar a produção relativa à utilização de fontes de energia renováveis e à cogeração [40]. A produção deve, dentro dos possíveis, respeitar a demanda, isto é, o consumo de eletricidade previsto para o mesmo período.

Ao longo de um dia, o consumo de energia elétrica apresenta, salvo exceção, uma previsibilidade sensivelmente elevada, pelo que, dadas as circunstâncias, devem ser colocadas em funcionamento os diversos tipos de produção disponíveis, de forma a garantir que o diagrama de carga seja respeitado e coberto. Desta forma, usando como caso de estudo Portugal, percebe-se que a produção deve acompanhar o diagrama do consumo, associado ao excesso de carga para bombagem (se necessário). Ora, caso a soma das diferentes produções seja suficiente para cobrir o diagrama total das cargas, não é necessária a importação, sendo até possível exportar, caso seja produzida em excesso.

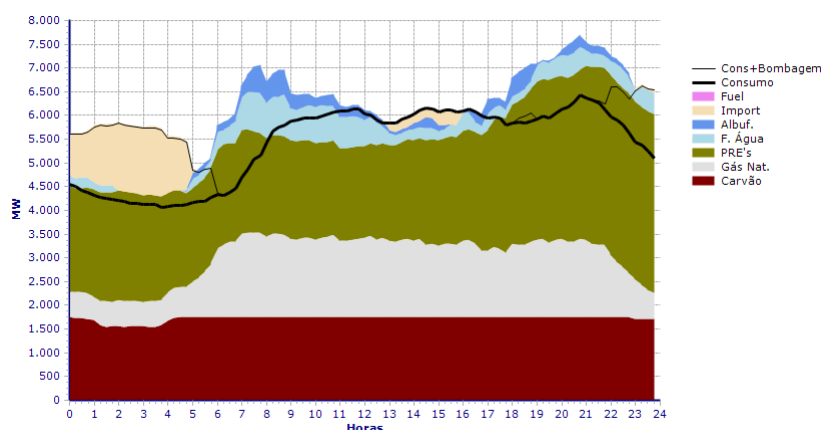


Figura 2.11: Estatística diária do SEE a 26 de abril de 2017 [5].

Como é possível verificar na Figura 2.11, temos um exemplo representativo das várias possibilidades a ocorrer durante um dia. Isto é, destaca-se inicialmente um valor razoável de importação, devido à obrigatoriedade de cobrir as necessidades de bombagem, contrariamente ao que acontece no final do dia, onde a produção por si só é suficiente para retificar a existência de bombagem. Ainda assim, durante o dia verificam-se períodos que garantem exportação, uma vez que a produção excede um pouco o requerido. Globalmente, o diagrama vai de encontro ao esperado, até porque se verificam as duas subidas de consumo habituais por volta das 12h00 e das 21h00.

De uma forma geral, é possível afirmar que, consoante o dia da semana (i.e., dia útil ou não), a curva do consumo tende a respeitar um padrão significativo. Ao longo do ano percebe-se que

existe uma pequena descida no consumo quando se aproximam as estações de maior calor (i.e., primavera e verão) e um aumento nas estações de maior frio (i.e., outono e inverno).

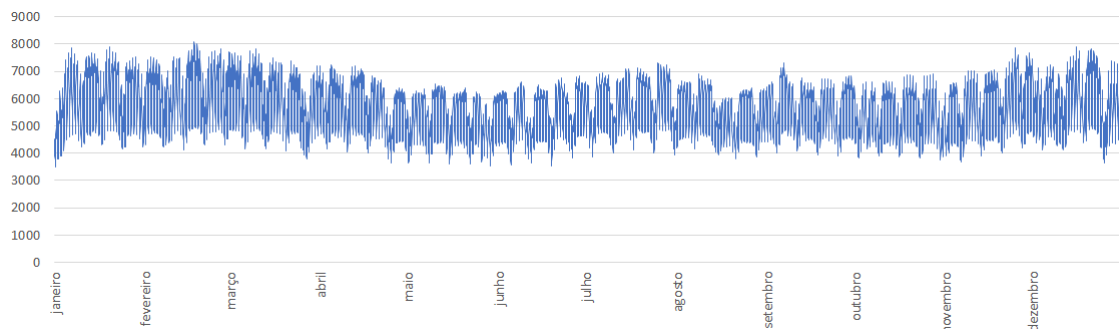


Figura 2.12: Estatística diária do consumo elétrico do ano de 2016 (*MWh*).

Como se constata pela Figura 2.12, num gráfico construído com base na informação disponível em [39], este comportamento ocorre de uma forma visível. Para além disso, 2016 foi um ano com um valor médio anual de 5613.3 *MWh*, a variar ao longo do tempo sensivelmente entre os 3000 *MWh* e os 8000 *MWh* em cada hora, sem grandes oscilações. Facilmente se entende que será possível estabelecer uma forte correlação entre os valores do consumo e algumas variáveis, o que permite obter valores previstos bastante razoáveis.

Capítulo 3

Ferramenta de previsão e procedimentos utilizados

Neste capítulo, para além duma exposição acerca das redes neuronais de regressão generalizada, pretende-se destacar os principais procedimentos e metodologia utilizada.

3.1 Rede Neuronal de Regressão Generalizada

3.1.1 Introdução

Um dos tipos de redes neuronais existentes é a rede neuronal de regressão generalizada (*GRNN*). Trata-se de uma rede baseada em memória que providencia a estimação de variáveis contínuas e converge para uma superfície de regressão subjacente. Ora, uma *GRNN* caracteriza-se por um algoritmo de aprendizagem de apenas uma passagem com uma estrutura massivamente paralela. Mesmo que os dados e a informação seja escassa, o algoritmo proporciona uma transição suave entre valores e pode ser usado para qualquer problema de regressão onde assumir linearidade não seja justificável. Tratando-se de uma rede paralela, é comum utilizar-se este tipo de rede em aplicações como a aprendizagem de modelos preditivos ou de controlo.

3.1.2 Abordagem e Estrutura

As *GRNN* são habitualmente usadas como funções de aproximação. Trata-se de um algoritmo de aprendizagem consideravelmente paralelo com o principal objetivo de obter um mapeamento perfeito entre um dado vetor de entradas e um vetor de saídas, com o mínimo erro. São redes que não requerem procedimentos de treino iterativo como as redes de retropropagação. No caso, aproximam uma qualquer função arbitrária entre os vetores de entrada e saída, e esboça uma função de estimação diretamente através do conjunto de treino. Ainda mais, é consistente afirmar que quanto maior o conjunto de treino, menor se torna, tendencialmente, o erro de estimação [41].

Trata-se de uma variante das redes neurais de base radial, que por sua vez são baseadas em redes de regressão. Ainda assim, mesmo assemelhando-se a uma rede radial, apresenta uma segunda camada um pouco distinta [42].

Uma *GRNN* consiste em quatro camadas [41]: a primeira uma camada de *inputs*; a segunda, uma camada onde cada neurónio representa um padrão de treino, fazendo com que também seja denominada por camada de padrão. Nesta camada, os neurónios escondidos calculam a distância *euclidiana* do ponto central do neurónio até ao caso de teste [43]; a terceira, uma camada linear ou camada de somatório; a última, a camada de *outputs*.

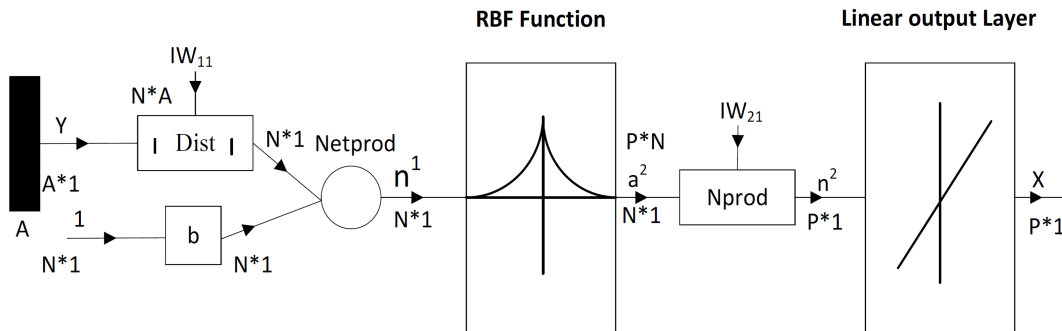


Figura 3.1: Estrutura Geral de uma *GRNN* [43].

Assim, percebe-se que, comparativamente a uma rede de base radial, apresenta um ligeira diferença na segunda camada (i.e., denominada *Special Linear Layer*). Note-se que, na segunda camada (i.e., *Radial Basis Layer*) existem tantos neurónios como pares de entrada/saída (e.g., se colocarmos como entrada 8760 valores por entrada, para 8760 valores para uma saída, resultará numa camada de base radial com 8760 neurónios) [44]. Aqui, o valor da entrada de cada neurónio desta camada é o valor pesado da distancia entre o vetor de *inputs* e o seu vetor de pesos. Ora, a rede de entradas de cada neurónio é o produto destas entradas pesadas e da sua *bias*. De notar que esta *bias* trata-se de um vetor de coluna composto por valores de $-\log(0.5)/Spread$ [44], onde *Spread* indica o parâmetro de suavização a estudar em 3.1.3. Por conseguinte, a saída de cada um destes neurónios é o resultado da passagem deste produto pela camada de base radial.

A camada seguinte, i.e., a terceira camada denominada *Special Linear Layer*, como referido, é um pouco distinta da original, das redes neurais de base radial. Ora, aqui, esta camada caracteriza-se por ser uma camada de somatório, onde são realizadas duas operações de soma dos valores anteriores. Finalmente, é realizada um mero quociente entre estes dois somatórios, da seguinte forma [45]:

$$\hat{Y}(X) = \frac{\sum_{i=1}^n Y^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \quad (3.1)$$

onde

$$D_i^2 = (X - X^i)^T (X - X^i) \quad (3.2)$$

De notar que σ representa o tal parâmetro de suavização. Percebe-se então que, ainda que as *GRNN* e as funções de base radial (*RBF*) tenham sido motivadas por diferentes princípios, a sua implementação é bastante similar. Ambas as redes têm um processamento de informação idêntico na camada escondida, para além de realizarem operações lineares na seguinte. Ora, a grande diferença é então que as *GRNN* executam uma média pesada enquanto que as *RBF* executam um somatório pesado. Decidir qual das duas será mais eficaz depende claro da sua aplicação e do tipo de problema a solucionar. Uma *GRNN* implica que se memorize grande parte da informação e, como tal, quando as restrições computacionais não são significantes, como em cenários onde é apenas necessário realizar uma previsão uma vez por dia, estas podem-se sobressair [46].

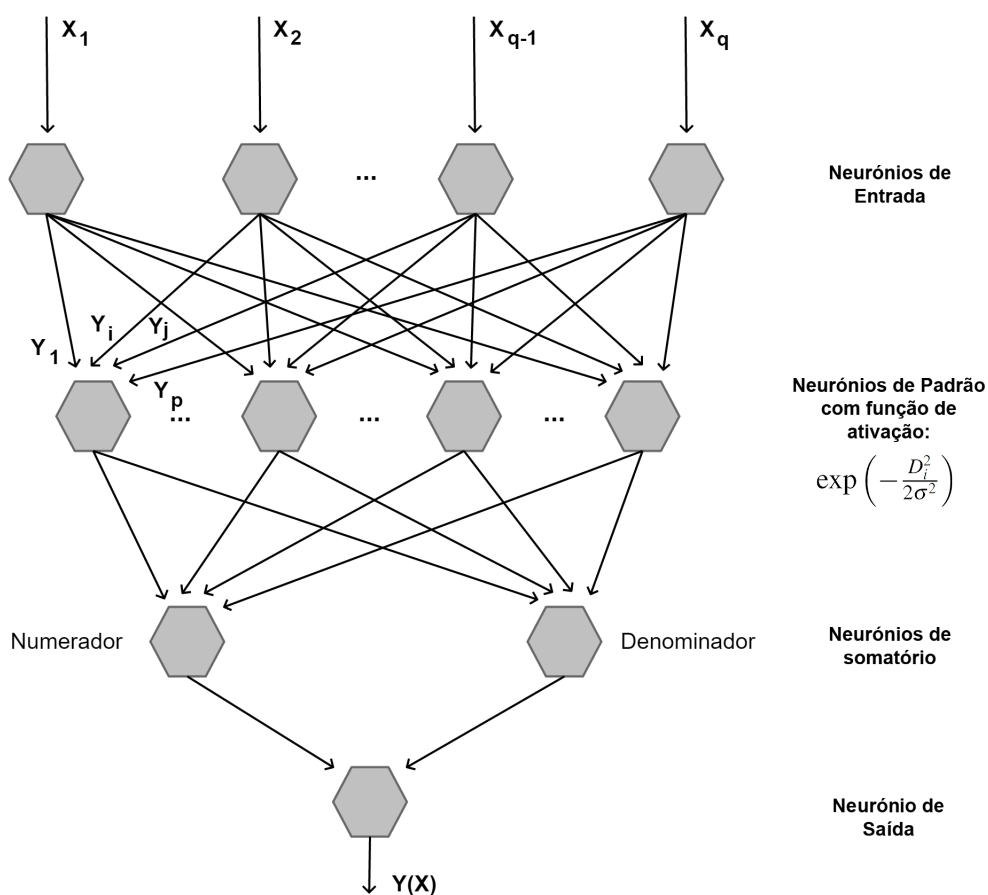


Figura 3.2: Arquitetura de uma *GRNN*

Assim, analisando a Figura 3.2, facilmente se estabelece uma relação da estrutura das *GRNN* com a sua arquitetura. Ora, aos neurónios de padrão chega um sinal Y_i que, juntamente com uma função de ativação (que por sua vez tem em consideração o parâmetro de suavização σ), resulta em dois novos sinais que atingem a camada de somatório em duas partes (i.e., numerador e denominador). Realizadas estas operações, apresenta-se um sinal de saída $Y(X)$.

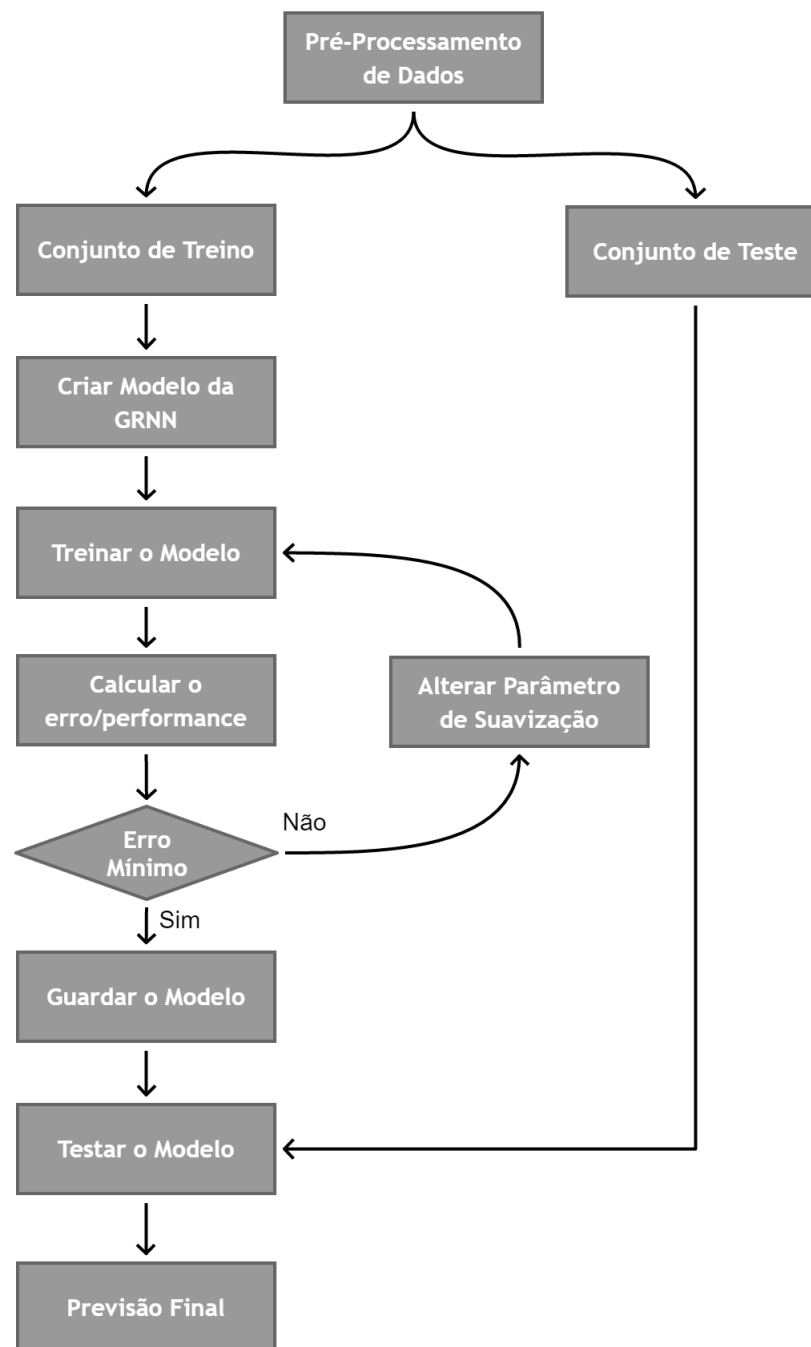
Perceba-se a importância de σ em todo o processo das *GRNN*. Valores diferentes deste parâmetro influenciarão claramente o padrão criado, e obviamente os resultados dos processos da camada em questão.

3.1.3 Normalização de Dados e Parâmetro de Suavização

Como procedimento de pré-processamento, é necessário escalar todas as variáveis de entrada para que todas pertençam à mesma gama de valores e variâncias. Esta necessidade parte do pressuposto de que a função de densidade de probabilidade subjacente deve ser estimada com um núcleo que tenha a mesma largura em cada dimensão. Ainda que não obrigatório, este procedimento é importante para conjuntos de dados finitos [45].

Este modelo de rede neuronal proposto por *Specht* representa uma aprendizagem através de um modelo simples, mas interpretável. Este requer a introdução de um parâmetro de suavização (Smoothing Parameter) que pode ser uniforme ou variável por dimensões. É importante referir este parâmetro, uma vez que surge como um importante fator no desempenho das *GRNN* [47]. Ainda que existam diferentes formas de obter um valor eficaz deste parâmetro, o método mais convencional passa por um processo de "tentativa e erro", onde é utilizado um estimador de erro (i.e., um indicador de performance) para cada iteração onde, é incrementado o valor de σ , até se deixar de notar um aumento na performance.

Então, é possível resumir um algoritmo de funcionamento de uma *GRNN* a:

Figura 3.3: Algoritmo de funcionamento de uma *GRNN*.

3.2 Obtenção e Tratamento de Informação

De forma a garantir conteúdo para a realização das previsões surge a necessidade de obter dados relativos ao preço da eletricidade e ao consumo de energia. Neste projeto estabeleceu-

se que iria ser tratada informação nacional¹ e como tal, impôs-se a necessidade da obtenção de diferentes tipos de informação, como:

- Valores horários por ano civil do preço de mercado da eletricidade em Portugal (€/MWh);
- Valores horários por ano civil da produção de energia em Portugal, por tipo de energia (MWh);
- Valores horários por ano civil da importação e exportação de energia em Portugal (MWh);
- Valores horários por ano civil do diagrama de consumo de energia em Portugal (MWh);
- Datas de feriados nacionais para os anos em estudo.

Acerca dos dados anuais e horários do preço de mercado da eletricidade em Portugal, utilizou-se a ferramenta da OMIE existente em [48], que tem disponíveis estes valores, com escala horária e por dia. Os valores extraídos estão representados em €/MWh.

Relativamente aos valores horários de produção e consumo, foi utilizada a ferramenta da REN disponível em [5], que disponibiliza esta informação também diariamente mas a cada 15 minutos. Importa referir que, relativamente à informação disponibilizada pela REN, a produção divide-se nos seguintes tipos:

- | | | |
|----------------|-----------------|---------------------|
| • Carvão; | • Fios de Água; | • PRE Eólica; |
| • Fuel; | • Bombagem; | • PRE Fotovoltaica; |
| • Gás Natural; | • PRE Hídrica; | • PRE Ondas. |
| • Albufeiras; | • PRE Térmica; | |

A juntar a estes valores de produção, consumo, importação/exportação, preço e feriados, temos as variáveis cíclicas habituais, i.e.:

- Mês;
- Dia da Semana (DDS);
- Hora.

3.2.1 Pré-processamento de Dados

Como referido em 3.1.3, de forma a garantir uma melhor uniformidade da informação, é necessário fazer passar o conjunto de entrada por um processo de normalização, isto é, um processo de ajuste de escalas. Ainda que também seja possível fazê-lo para o conjunto de saída, é de certa

¹Dada a complexidade do sistema a analisar, e disponível abertamente a informação requerida, foram analisados os anos de 2011 a 2016, inclusive.

forma indiferente já que, aquando da comparação com os valores reais, seria necessário realizar uma desnormalização, tornando o processo contraproducente.

Existem diversos métodos disponíveis, tais como o método *Min-Max*, o método *Z-Score* ou o método *Decimal Scalling*. O método escolhido para este estudo foi o método *Z-Score*, onde é necessário calcular a média e o desvio padrão do conjunto a normalizar.

Aqui, o valor x do conjunto A é então escalado para x' onde \bar{A} e σ_A correspondem à média e ao desvio padrão do conjunto de entrada A [49]:

$$x' = \frac{x - \bar{A}}{\sigma_A} \quad (3.3)$$

Para as variáveis cíclicas o tratamento é distinto. Assim, as variáveis de entrada (x) *Mês*, *DDS* e *Hora* são, cada uma delas, divididas em duas novas entradas (x_1 e x_2):

$$x_1 = \sin\left(\frac{2\pi x}{\rho_x}\right) \quad (3.4)$$

$$x_2 = \cos\left(\frac{2\pi x}{\rho_x}\right) \quad (3.5)$$

onde x representa o valor original e ρ_x representa o número total de possibilidades para a variável (i.e., 24 para a variável *Hora* ou 12 para a variável *Mês*).

De referir que, pretendendo-se realizar uma previsão de dados, para o dia t , é necessário impor novas variáveis para cada uma das entradas não sazonais (*Produção*, *Consumo*, etc.) resultando um conjunto de novas variáveis deslocadas. Falamos então, no caso para a previsão das variáveis correspondentes a um horizonte temporal $t - h$ (h horas antes):

$$x_h(t) = x(t - h) \quad (3.6)$$

Tomando como exemplo os dados anuais e horários de uma determinada variável (e.g., 8760 exemplos), caso se pretenda executar os procedimentos associados à expressão 3.6, os dados iniciais associados ao maior desvio (e.g., 168 horas ou exemplos), vão ser excluídos do conjunto de exemplos total, de forma a garantir a existência de exemplos de todas as variáveis para todas as horas.

3.2.2 Feriados

Relativamente à variável de entrada referente aos feriados, surgem diversas possibilidades de tratamento de dados. Ora, existe a possibilidade de tratar estes dia como se de um domingo se tratasse, devido ao comportamento das variáveis ser de certa forma semelhante entre ambos. Ainda assim, é aconselhável impor um tratamento singular, adicionando uma nova variável binária que indica se se trata, ou não, de um feriado. Também como no caso das variáveis não cíclicas, a

variável feriado deve ser sujeita a uma normalização pelo método *Z-Score* (ver expressão 3.3) e deve também ser separada numa ou mais variáveis, como descrito na expressão 3.6.

3.2.3 Indicadores de Desempenho

É importante avaliar o desempenho de uma previsão usando para isso previsões genuínas e completas. Isto é, não é coerente verificar o quão bem um modelo se encaixa nos dados passados ou históricos. A precisão de uma previsão apenas pode ser determinada quando se considera o quão bem desempenha o seu papel em novos dados, que não foram usados para estimar o modelo.

Na realização de previsões, o conjunto total é dividido em conjunto de treino e conjunto de teste com este último, habitualmente, a ser pelo menos tão grande quanto o tamanho do horizonte a prever.

De uma forma geral, relativamente ao desempenho de um modelo, importa notar [50]:

- Um modelo pode encaixar bem na informação de treino mas não é necessariamente obrigatório que realize boas previsões;
- Uma boa correlação pode ser alcançada usando parâmetros suficientes;
- Sobre-treinar um modelo para um conjunto de dados específico não é recomendável, pela possibilidade de falha na identificação padrões sistemáticos.

Segundo [51], existem diversos métodos de medição da precisão preditiva, entre os quais o Mean Average Error (*MAE*) e o Root Mean Square Error (*RMSE*). Estes podem ser usados para avaliar a precisão de previsões *ex-post* e *ex-ante*. Uma previsão *ex-post* é aquela na qual são usados os valor verdadeiros das variáveis exógenas e uma previsão *ex-ante*, ainda que não muito interessante, é aquela onde são usados valores aproximados destas mesmas variáveis.

Para além disto, segundo [52], é possível distinguir os métodos de avaliação em métodos autónomos ou independentes e métodos de precisão relativa. Os primeiros são aqueles que podem ser obtidos sem a referência a previsões adicionais, estão normalmente associados a funções de precisão quadráticas ou baseadas em mínimos absolutos. Os métodos baseados nos valores mínimos quadráticos, têm vindo a ser selecionados como um dos mais eficazes para avaliação de previsão de séries individuais.

Supondo então que o nosso conjunto de dados é denotado por y_1, \dots, y_T , e que o dividimos em duas secções: conjunto de treino (y_1, \dots, y_N) e conjunto de teste (y_{N+1}, \dots, y_T). De forma a verificar a precisão de um modelo de previsão, deve-se estimar os parâmetros utilizando o conjunto de treino e prever os próximos $T - N$ exemplos. Após isto, estes valores previstos vão ser sujeitos a uma comparação com os valores de teste.

Uma previsão com um passo h , pode ser descrita como $\hat{y}_{N+h|N}$, com \hat{y} a representar um valor previsto e não o valor de exemplo e o índice a indicar que estamos a estimar y_{N+h} utilizando todos os exemplos até ao tempo N , inclusive.

Os erros de previsão são a diferença entre os valores reais no conjunto de teste e aqueles previstos. Daí,

$$e_t = y_t - \hat{y}_{t|N}, \quad t = N + 1, \dots, T. \quad (3.7)$$

Com isto, e relativamente aos indicadores dependentes de escala (i.e., indicadores que se encontram na mesma escala e grandeza que os dados [53]), os mais comuns são o *MAE* e o *RMSE* [51]. Acerca dos métodos independentes da escala, ou métodos percentuais, um dos mais comuns é o *MAPE*.

MSE Mean Squared Error, é também usado em vários algoritmos de treino no processo iterativo, ou permite a decisão relativamente ao critério de paragem

$$MSE = \frac{1}{T} \sum_{i=1}^T e_i^2 \quad (3.8)$$

RMSE Root Mean Squared Error, pode ser usado como medida relativa para comparar previsões da mesma série de valores, mas de diferentes modelos, quanto menor o erro, melhor capacidade de previsão terá o modelo.

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T e_i^2} \quad (3.9)$$

Um problema associado a este indicador passa pela sua variância de previsão ao longo do tempo. Pode variar devido à não-linearidade do modelo e devido à variação das variáveis exógenas (se existentes). Segundo [51], não é possível fazer uma interpretação estatística rigorosa do *RMSE* porque esta não é uma estimação de nenhum parâmetro do modelo.

MAE Mean Absolute Error, também é dependente da escala mas é menos sensível a largos desvios do que os métodos quadráticos convencionais.

$$MAE = \frac{1}{T} \sum_{i=1}^T |e_i| \quad (3.10)$$

MAPE Mean Absolute Percentage Error, é um indicador percentual que tem a vantagem de ser independente da escala e é frequentemente utilizado para comparar desempenhos entre diferentes conjuntos de dados.

$$MAPE = \frac{1}{T} \sum_{i=1}^T |p_i| \quad (3.11)$$

com

$$p_t = 100 \cdot \frac{e_t}{y_t} \quad (3.12)$$

Este indicador faz uma medição baseada nos erros percentuais, ainda que com a desvantagem de tomar valores infinitos ou indefinidos quando $y_t = 0$ ou valores muito elevados quando y_t é muito perto de zero.

Outro problema com erros percentuais que é muitas vezes mencionado passa por estes assumirem uma escala baseada em quantidade. Se y_t for medido por exemplo numa moeda, quilos ou noutra qualquer quantidade, faz sentido usar percentagens. Porém, utilizar um erro percentual deixa de fazer sentido quando se pretende medir a precisão da previsão de temperatura nas escalas *Celsius* ou *Fahrenheit* porque estas não medem quantidades. Entenda-se que a principal questão a colocar acerca do *MAPE* prende-se com a sua forte dependência do valor do denominador, resultando em problemas quando o seu valor for próximo de zero, ou mesmo zero. Ora, visto de outro prisma, apenas faz sentido utilizar erros percentuais se mudar a escala não interferir na percentagem. Alterar y_t de quilos para libras (*pounds*) resultará na mesma percentagem, não acontecendo o mesmo caso se altere y_t de Celsius para Fahrenheit.

MAPE' Este indicador, uma variante do *MAPE*, tem como função filtrar os casos anómalos da medição percentual acima referidos. Neste caso, faz uso da medição do *MAE*, juntamente da média do conjunto, de forma a garantir que não existam valores finais infinitos ou indeterminados.

$$MAPE' = \frac{MAE}{\frac{1}{T} \sum_{i=1}^T y_t} \quad (3.13)$$

Capítulo 4

Metodologia Utilizada

4.1 Introdução

Neste capítulo faz-se uma consolidação das metodologias utilizadas para a realização das simulações, assim como outros procedimentos ou tomadas de decisão importantes no desenrolar do projeto. Encontra-se assim dividido nas seguintes secções:

- Software e ferramentas utilizadas para previsão;
- Aplicação desenvolvida para realização de simulações;
- Escolha do melhor modelo para previsões.

4.2 *Software* Utilizado

4.2.1 Organização e Tratamento de Dados

Relativamente à organização e processamento de informação optou-se por manter os ficheiros de dados em folhas de cálculo associadas ao *software* Microsoft® Excel. Então, para cada conjunto de dados existentes (i.e., anos 2011 a 2016) surge um ficheiro que contém uma página associada aos valores não processados (utilizados para as variáveis de saída) e ainda uma folha associada às variáveis processadas (i.e., normalizadas pelos métodos descritos em 3.2.1). Ainda assim, de forma a garantir uma celeridade no processo de importação e tratamento de dados, utilizam-se estes valores originais em ficheiros *.mat* (i.e., para utilização no *software* [54]) que serão tratados e normalizados de forma automática.

4.2.2 Execução de Simulações e Cálculo de Indicadores de Performance

Ainda que disponíveis diversas ferramentas e *softwares* para a realização de previsão, optou-se por utilizar o *software* Mathworks MATLAB R2016b para a realização de todas as operações subsequentes ao tratamento de dados.

A execução de um elevado número de simulações, maioritariamente para conjuntos de informação e quantidade de variáveis elevadas torna o processo de treino do algoritmo em estudo (ver Secção 2.5.5) moroso. Ainda mais, de forma a obter uma melhor visualização dos indicadores e das curvas originais e previstas, e de forma a otimizar todo o processo de testes, optou-se pelo desenvolvimento de uma aplicação, a analisar em 4.3.

4.2.3 Ferramenta para Criação e Treino de GRNN

O *software* [54] contém uma ferramenta independente para criação e treino de uma GRNN – *newgrnn* – que requer apenas três entradas, uma matriz de entrada, uma matriz de saída, e o valor do parâmetro de suavização:

$$net = (X, Y, spread) \quad (4.1)$$

onde X e Y representam as matrizes de valores de entrada e de saída, respetivamente e onde *spread* representa o valor do parâmetro de suavização.

De notar que, na camada de base radial (ver 3.1.2) a *bias* toma o valor de $\frac{\sqrt{-\log 0.5}}{spread}$ pelo que o valor a atribuir ao parâmetro de suavização terá impacto direto no desempenho da rede.

4.2.4 Ferramenta para Criação e Treino de Feed-Forward ANN

Acerca da implementação e treino de redes neuronais convencionais *Feed-Forward*, o *software* [54] apresenta diversas soluções relativamente à sua estrutura e algoritmo de treino. Ainda assim, pela sua maior rapidez de treino, faz-se uso da ferramenta *nftool* e assim, do algoritmo de treino *trainlm* – *Levenberg-Marquardt* (ver Secção 2.5.5).

Desta forma, resulta uma rede *Feed-Forward* de duas camadas, com os neurónios da camada escondida com funções de ativação do tipo sigmoide, sendo a segunda constituída por neurónios lineares. A rede é treinada pelo método *Levenberg-Marquardt*, que reflete um algoritmo de retro-propagação.

O processo é iterativo, e faz uso do indicador *MSE* para o cálculo do erro por iteração. Se ao fim de seis iterações não se verificar uma redução do erro, dá-se por terminada a simulação, com resultados finais.

4.3 Aplicação Desenvolvida para Realização de Simulações

A previsão do preço da eletricidade e do consumo de energia obriga a garantir uma robustez de resultados notória, assumindo-se a necessidade da realização de um número elevado de simulações (i.e., para escolha do melhor modelo, afinação, etc.). O processo de migração manual de dados de uma folha de cálculo para uma outra aplicação é sujeito a falhas e obriga a execução constante das ferramentas de ajuda. Devido à existência de duas ferramentas independentes para ambos os tipos de redes neuronais a estudar, é possível programar execução de ambas em série, com auxílio

computacional na obtenção de alguns parâmetros essenciais para a otimização de determinados métodos.

Sendo ainda necessário o cálculo de diferentes indicadores, verificação de tempos de execução, e avaliação outros parâmetros, a significância do desenvolvimento de uma aplicação é evidenciada pela forte elasticidade da linguagem do software em utilização [54].

Perante isto, desenvolveu-se uma aplicação capaz de criar um comitê de cenários de escolha e organização de dados a trabalhar, simulações múltiplas dos métodos em estudo, cálculo de indicadores de performance e visualização de curvas. Realizada na linguagem nativa do *software* [54], está então preparada para:

1. Seleção da variável a prever;
2. Seleção das variáveis de entrada;
3. Seleção dos horizontes temporais a utilizar;
4. Seleção do Ano a estudar;
5. Seleção do método de amostragem de dados;
6. Seleção dos pesos de treino e teste;
7. Visualização dos indicadores de performance, por método;
8. Visualização dos tempos de execução, por método;
9. Visualização de curvas comparativas entre valores reais e previstos, por método.

Não querendo limitar a execução das simulações e consequente visualização de resultados a um ambiente mais pálido, como a linha de comandos, optou-se por rematar a nossa aplicação com implementação de uma simples interface gráfica.



Figura 4.1: Interface gráfica da aplicação desenvolvida.

De forma a garantir o perfeito funcionamento desta ferramenta, há apenas que respeitar o preenchimento do modelo base das folhas de cálculo, onde devem ser presentes os dados obtidos e descritos em 3.2.

4.3.1 Pré-processamento de Dados

Como referido em 4.2.1 os dados relativos a cada um dos anos disponíveis (i.e., 2011 a 2016), estão disponíveis de forma independente em ficheiros *.mat*. Assim, após ser carregada a informação relativa ao ano desejado, é realizado o tratamento de dados relativo à implementação dos horizontes temporais impostos e consequente normalização de dados.

4.3.2 Amostragem de Dados

O processo de seleção de dados é uma etapa crucial para o desenvolvimento de toda a simulação. Nesta secção, faz-se um levantamento dos diferentes algoritmos implementados para divisão dos dados, para além da sua divisão em conjunto de treino e teste.

A desagregação dos exemplos em dois novos conjuntos é uma das etapas nucleares do processo de treino. É habitual considerarem-se valores na ordem dos 75% para o conjunto de treino e o restante para o conjunto de teste e validação (se existente). O valor por defeito colocado na aplicação aplica um peso de 75% para o treino e os restantes 25% para teste.

Ainda assim, surge a possibilidade do utilizador optar por escolher pesos específicos (*Specify Weights*) ou por deixar a aplicação realizar uma seleção de pesos aleatória, dentro de certos limites. Preante isto, é escolhido um valor aleatório entre 70% e 75% para o conjunto de treino, resultando um peso de treino W_{train} . Escolhido o peso de treino, a percentagem remanescente é então atribuída ao peso de teste, W_{test} . Assim, a ferramenta está preparada para fazer uma seleção aleatória de pesos da seguinte forma:

$$W_{train} = random(train), \quad 0.70 \leq train \leq 0.75 \quad (4.2)$$

$$W_{test} = 1 - W_{train} \quad (4.3)$$

resultando,

$$T_{train} = T \cdot W_{train} \quad (4.4)$$

$$T_{test} = T \cdot W_{test} \quad (4.5)$$

onde T representa a quantidade total de exemplos, T_{train} a quantidade de exemplos para o conjunto de treino e T_{teste} a quantidade de exemplos para o conjunto de teste.

Ainda que para as *GRNN* a seleção seja simples de executar e compreender, para o caso das *ANN* convencionais, treinadas com a ferramenta descrita em 4.2.4, já existe uma divisão de dados incorporada. Ora, assumindo que etapa de teste de dados previstos é feita de uma forma independente da ferramenta incorporada do software [54], a preparação da montagem da *ANN* deve ser feita de forma distinta do habitual, não sendo necessário reservar parte dos exemplos para teste (procedimento a analisar em 4.3.4).

A ferramenta está preparada para permitir que o utilizador opte entre três diferentes modelos de amostragem, isto é, permite uma seleção de dados aleatória, ordenada ou misturada:

Seleção Aleatória A seleção aleatória, tal como nome indica, permite, partindo de um dado conjunto de exemplos, realizar um sorteio aleatório mas emparelhado entre exemplos. Ora, calculados os totais de exemplos para cada conjunto, são escolhidos T_{train} exemplos de forma aleatória e desordenada de forma a constituir o conjunto de treino. Mapeados os índices dos exemplos escolhidos para o conjunto de treino, são então selecionados aqueles ainda não escolhidos para o conjunto de teste, também de forma desordenada.

De notar que, como a amostragem de dados de ambos os conjuntos foi monitorizada com um mapeamento das seleções, de forma paralela à amostragem dos exemplos do nosso conjunto de dados originais (output), é realizada uma seleção no conjunto de dados normalizados (input), garantindo um emparelhamento fiel da seleção executada (ver Figura 4.2).

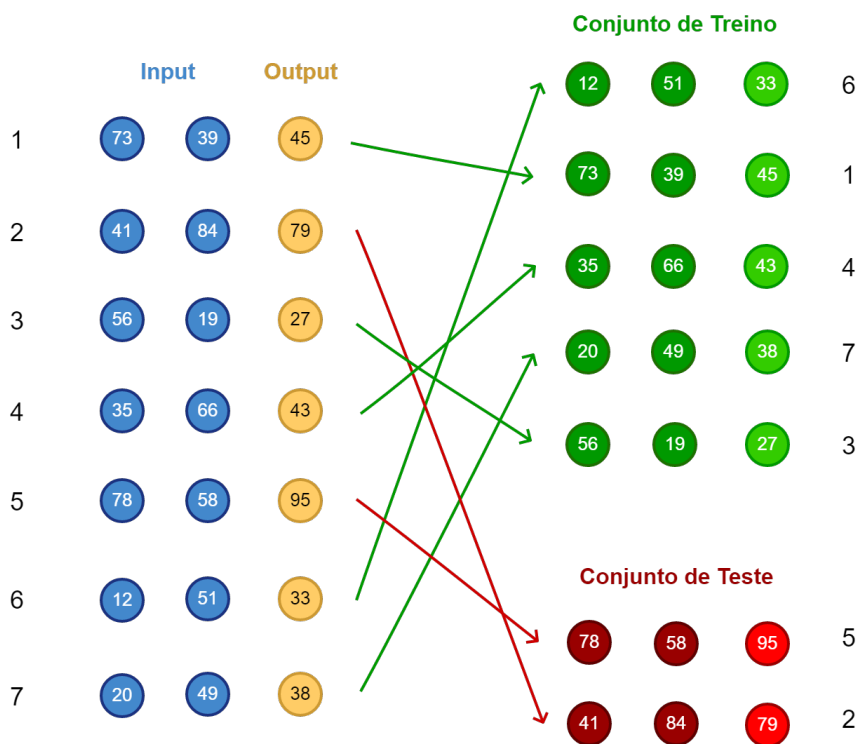


Figura 4.2: Exemplo de uma amostragem aleatória.

Ainda que este algoritmo de seleção de dados administre de uma forma firme a informação a trabalhar, num caso prático torna-se inviável a sua execução. Entenda-se que, garantida uma grande elasticidade de dados selecionados, os resultados de previsão podem ser mais robustos. Ainda que não seja exequível numa previsão real, o método é muito útil na perceção e teste do melhor modelo a utilizar (i.e., escolha das variáveis de entrada).

Seleção Ordenada Este modelo de seleção de dados, garante uma separação direta e ordenada dos exemplos existentes, consoante os pesos de treino e teste. Ora, para o conjunto de treino são escolhidos os primeiros T_{train} exemplos de forma ordenada com os restantes T_{test} exemplos a constituírem o conjunto de teste (ver Figura 4.3).

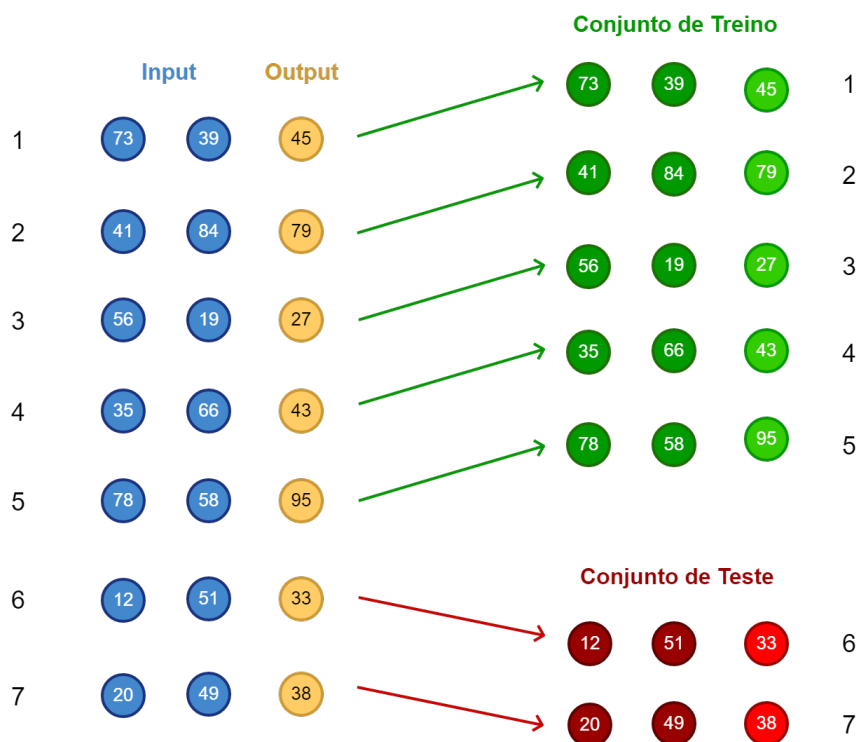


Figura 4.3: Exemplo de uma amostragem ordenada.

Ao contrário do que acontecia no método aleatório, o algoritmo ordenado encaixa na perfeição naquilo que será uma previsão real. Isto é, temos um dado conjunto de exemplos históricos e queremos prever os seguintes.

Seleção de Semanas Misturadas O algoritmo de semanas misturadas destaca-se por ser uma mistura dos anteriores, ou seja, garante continuidade de valores mas ao mesmo tempo tem a mais valia de conseguir extrair exemplos de sensivelmente todas as zonas do conjunto de dados. Ora, o algoritmo está preparado, sem auxílio dos pesos, extrair sensivelmente $\frac{4}{5}$

dos exemplos para o treino e os restantes $\frac{1}{5}$ para o teste. Isto ocorre porque, de uma forma cíclica e ordenada, são seleccionadas quatro semanas de dados para treino e uma para teste (ver Figura 4.4).

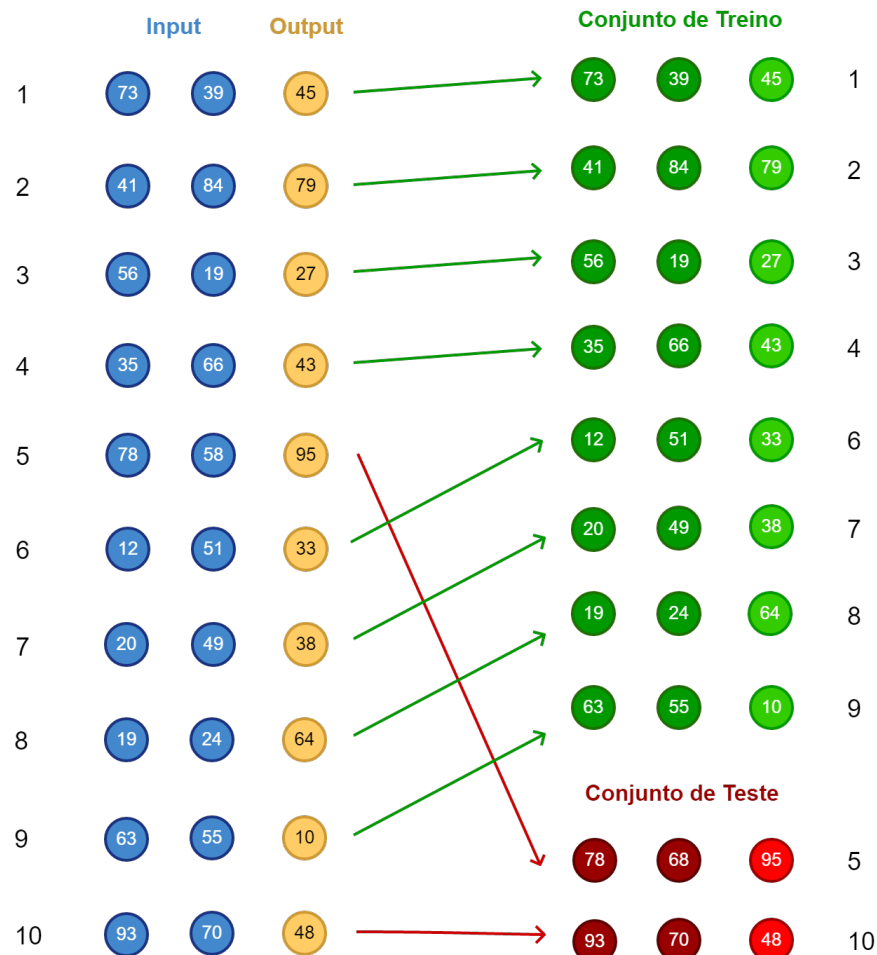


Figura 4.4: Exemplo de uma amostragem de semanas misturadas.

Este método também não é fiel ao estudo de casos reais e como tal, é um método capaz de auxiliar a sustentar a escolha dos modelos finais de previsão.

Seleção Adaptativa (hora-a-hora) De forma a querer impor um modelo cabal mas fortemente associado a uma situação real, optou-se por simular um algoritmo adaptativo. Este, para o período em questão, realiza uma previsão para um exemplo, e não para um conjunto de exemplos. Ora, caso queiramos prever 24 horas de um dia, o algoritmo irá utilizar o conjunto de informação fornecido até à hora em questão para prever a primeira hora desse conjunto de 24 horas. Prevista esta hora, e requerida a previsão da seguinte, a anterior é associada ao

conjunto de treino para a estimação da nova hora. O processo repete-se ciclicamente até ser previsto todo o conjunto requerido (ver Figura 4.5). Em suma, trata-se de um algoritmo que permite fazer a previsão para a hora seguinte.

Ainda que preparado para realizar a previsão exemplo a exemplo (i.e., hora a hora), como se realiza a previsão via *GRNN* e *ANN*, existe necessidade rever o processo. Para o caso das *GRNN*, ainda que se deseje prever um grande conjunto de exemplos, a nível computacional todo o progresso é razoavelmente célere. Porém, para o caso da previsão via *ANN*, a aplicação de uma rede neuronal com algoritmo de retropropagação, para cada um dos exemplos, pode tornar o processo demasiado lento, isto caso o número de exemplos a prever seja considerável (e.g., previsão de 3 meses ou 2160 horas, aproximadamente). Portanto, caso o período a prever seja substancialmente longo, é possível a alteração da escala de previsão por exemplo, de escala horária para diária. Assim, caso seja requerida a previsão direta de um dia completo e não contínua (i.e., hora a hora), a seleção de dados mantém a mesma estrutura, mas desta feita, por blocos de 24 exemplos (i.e., 24 em 24 horas).

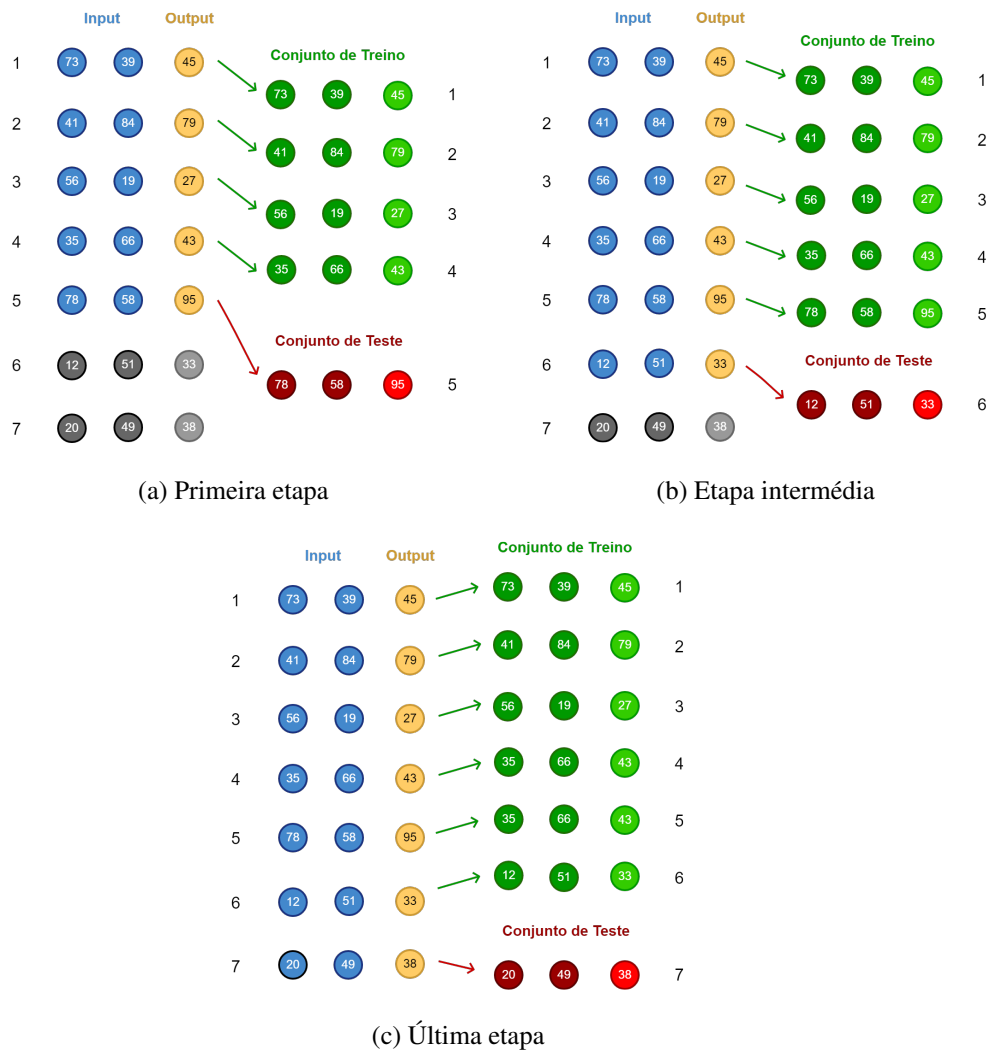


Figura 4.5: Exemplo de uma amostragem adaptativa hora-a-hora.

Ainda que o processo se torne massivamente pesado quando em comparação com o totalmente ordenado, garante uma maior presença da informação para cada previsão e um aumento da precisão, quando em comparação com os anteriores. De notar que, se necessário, existiria ainda a possibilidade de alterar o algoritmo de seleção de dados de forma a que nem todos os exemplos anteriores fossem selecionados, deixando progressivamente de forma os mais antigos, a cada iteração.

Seleção Adaptativa (dia-a-dia) Ainda que o método de seleção adaptativa hora-a-hora seja benéfico, para a previsão do preço da eletricidade e do consumo de energia elétrica, aquilo que é o requisito padrão aponta para uma previsão para o dia seguinte querendo-se a informação para todo o dia seguinte naquele exato momento. Ora, como aquele caso requeria um conhecimento constante dos resultados da hora anterior, não é exequível para a previsão de um bloco de 24 horas independentes. Desta forma, implementa-se um modelo que, de forma análoga ao anterior, realiza um deslizamento de dados para constituição do conjunto de treino. Ainda que, neste caso, a previsão das 24 horas (i.e., de h a $h + 23$) é realizada em bloco (apenas com o treino de uma rede neuronal) e apenas com os dados conhecidos até ao início deste mesmo dia (i.e., até $h - 1$). Previstos os dados do primeiro dia em questão, para a previsão do dia seguinte já são adicionados ao conjunto de treino os valores reais do dia anterior, e assim sucessivamente, de forma iterativa e cíclica.

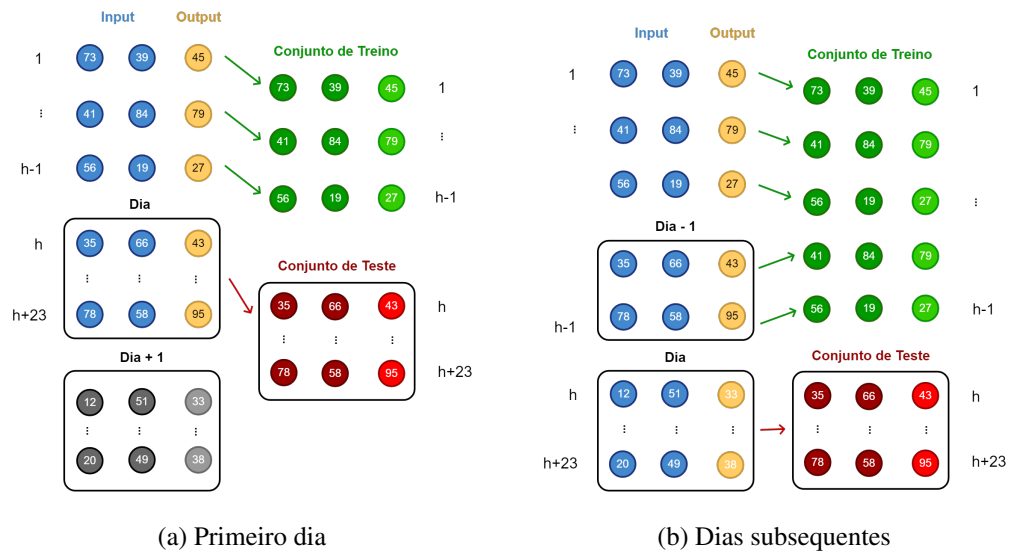


Figura 4.6: Exemplo de uma amostragem adaptativa dia-a-dia.

Note-se que, para o caso das *GRNN*, como é treinada uma rede por cada dia a prever, a estimação do parâmetro de suavização é realizada a cada interação (i.e., dia-a-dia). Desta forma, ao invés de se impor um valor inalterável, o algoritmo está preparado para adaptar a suavização de acordo com os novos dados incluídos, de forma progressiva. No caso das *ANN*, também é realizado o procedimento estudado em 4.3.4. De entender que, este modelo

é aquele que se aproxima mais de uma situação real para previsão diária e é fortemente motivado pela rapidez de treino das *GRNN*, quando comparadas com as *ANN*.

4.3.3 Previsão através de *GRNN*

O processo de previsão através de *GRNN* desenvolvido para a aplicação em questão é bastante simples. Contrariamente ao que acontece no caso das *ANN*, na sua aplicação computacional desenvolvida (ver Secção 4.3.4), onde é repetida a execução de treinos, no caso das *GRNN*, a execução é singular. Graças ao seu perfil unidirecional, a repetição de treinos é inconsequente, apresentando sempre os mesmos resultados, para um mesmo conjunto de dados. Ora, dado que o treino das *GRNN* é um processo determinístico (ao contrário das *ANN*), o resultado do treino é sempre o mesmo, não se justificando neste caso a repetição de treinos e a consequente seleção dos melhores modelos.

Ora, para o caso da previsão via *GRNN*, a aplicação está preparada para simplesmente treinar uma rede e fazer uso desta para prever os valores necessários através das ferramentas do *software* [54], e como descrito em 4.2.3.

Ainda assim, como estudado em 3.1.3, a escolha do parâmetro de suavização é crucial. Como tal, e de forma a seguir as indicações, a parametrização desta componente é realizada de forma iterativa, até que se obtenham os melhores resultados. Isto é:

1. Atribui-se um valor inicial e perto de zero a este parâmetro (e.g., *spread* = 0.02);
2. Realiza-se uma separação temporária do conjunto de treino original num conjunto de treino temporário e num conjunto de validação, para que se consiga compreender o impacto da alteração deste parâmetro. Ora, são seleccionados de forma aleatória 75% e 25% do total dos exemplos do conjunto de treino original para constituírem os conjuntos de treino temporário e validação.
3. Realiza-se uma previsão com este valor;
4. Usa-se um indicador de desempenho para verificar o erro associado (e.g., *MAE*) entre o conjunto de validação e o conjunto previsto;
5. Aumenta-se progressivamente o valor do parâmetro e repetem-se os dois pontos anteriores até que o erro aumente;
6. Verificado um aumento do erro, estabelece-se o parâmetro de suavização como sendo o valor encontrado antes deste acréscimo.

Alcançado o parâmetro de suavização ideal, é treinada uma rede com o total dos exemplos do conjunto de treino.

4.3.4 Previsão através de *Feed-Forward ANN*

Relativamente ao treino de redes neuronais convencionais, no caso com o algoritmo de treino *Levenberg-Marquardt*, o procedimento é relativamente diferente. Ora, excetuando no caso de

amostragem adaptativa, a processo de previsão via *ANN* convencionais segue as seguintes diretrizes:

1. **Estruturação do modelo de rede.** Aqui, estabelecem-se os seguintes parâmetros:

- Número de neurónios da camada escondida igual a:

$$hiddenLayerSize = 3 \cdot variables + 2$$

com *hiddenLayerSize* a representar o número de neurónios e *variables* a representar o número de variáveis de entradas.

- Pesos dos conjuntos de treino, validação e teste. Como referido, aqui retira-se praticamente todo o peso à parcela de teste, uma vez que existe um conjunto de exemplos especificado para o teste, de forma independente à incorporada no modelo das *ANN*. Assim, para os conjuntos de treino, validação e teste, respetivamente:

$$net.divideParam.trainRatio = 70/100$$

$$net.divideParam.valRatio = 29/100$$

$$net.divideParam.testRatio = 1/100$$

2. **Realização das previsões.** Para a realização das previsões (excetuando em modelos adaptativos), respeita-se um procedimento previamente estabelecido:

- Previsão através de dez redes neuronais, independentes entre si;
- Cálculo da média aritmética dos valores previstos, por cada exemplo (i.e., por hora), resultando num único conjunto de previsões;

4.3.5 Apresentação de Resultados

Relativamente à apresentação de resultados, a aplicação está preparada para disponibilizar ao utilizador os valores dos indicadores de desempenho, assim como uma comparação entre curvas reais e previstas.

Obtidos os valores previstos para ambos os modelos a aplicação executa o cálculo dos indicadores de desempenho descritos em 3.2.3. Também como referido em 3.2.3, para o caso de conjuntos de dados cujos exemplos possam atingir valores iguais ou muito próximos a zero, não se deve considerar o indicador *MAPE*. Ora, ainda que calculável, não se considera este indicador na previsão do preço da eletricidade.

De forma a garantir uma visualização gráfica dos resultados obtidos, são então apresentados dois gráficos comparativos entre os resultados da previsão via *GRNN*, *ANN* e o conjunto de teste (i.e., valores reais).

Capítulo 5

Apresentação e Análise de Resultados

5.1 Escolha do Melhor Modelo para Previsões

De forma a escolher um modelo robusto e que garanta um equilíbrio da realização dos testes, surge a necessidade de pressupor valores de referência para as previsões a realizar. Como tal, como modelo base para a previsão do preço da eletricidade e do consumo de energia elétrica, estabelece-se como ponto de partida, os resultados obtidos em previsões simples através da análise de séries temporais *Box-Jenkins* (ver secção 2.4). Daqui serão, ou não, adicionadas novas variáveis, de forma a impor um modelo de variáveis de entrada que seja sólido e que não introduza incerteza nos resultados, tal como pode acontecer no caso das ANN, em modelos com poucas variáveis de entrada. Ora, ainda que um modelo com apenas duas ou três variáveis de entrada possa, nalgumas tentativas resultar num erro mais reduzido, caso se entenda que a experiência é relativamente volátil, dá-se preferência a um modelo devidamente preparado para a execução de diversas simulações, e que apresente resultados com poucas flutuações.

5.1.1 Preço da Eletricidade

De forma a que a comparação das simulações seja efetuada de uma forma equivalente, é necessário estabelecer um modelo de previsão de base. Ora, analisando de uma forma célere a correlação entre a variável a prever (i.e., o preço da eletricidade) e as variáveis de entrada disponíveis (ver Secção 3.2), realizam-se simulações sucessivas, com a adição e/ou remoção de variáveis de entrada, até que se obtenha o modelo ideal.

No caso da previsão do preço da eletricidade, habitualmente para o dia seguinte, de forma a garantir a possibilidade da realização desta previsão com a antecedência requerida, opta-se por utilizar horizontes temporais de atraso no mínimo com quarenta e oito horas de atraso (i.e., dois dias). Desta forma, como se introduzem dados referentes às 48 horas anteriores, é possível realizar a previsão com a antecedência estipulada em mercado.

5.1.1.1 Preço da eletricidade em 2016

Usando como exemplo o ano de 2016, e de forma a realizar uma previsão singela via *Box-Jenkins*, é necessário, antes disso, analisar o ACF e a PACF, nunca antes de sustentar ideias na verificar dos valores da curva do preço ao longo do ano.

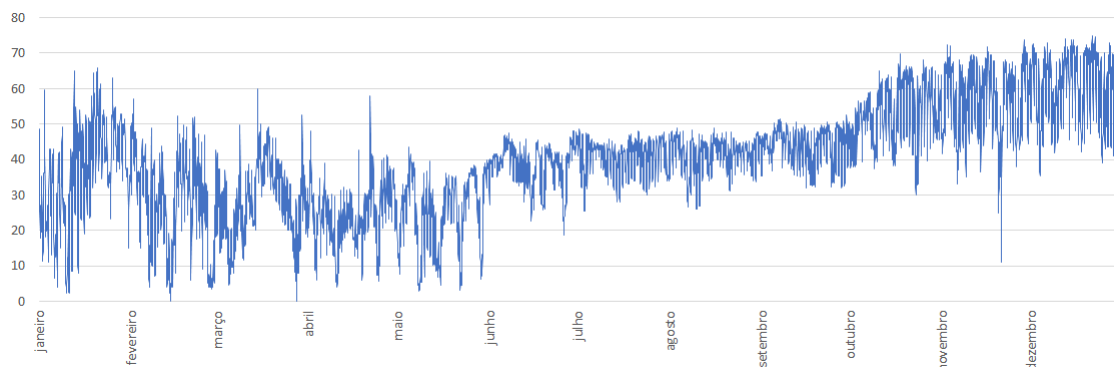


Figura 5.1: Evolução do preço da eletricidade em 2016 (€/MWh).

Como é possível verificar na Figura 5.1, o preço da eletricidade ao longo do ano de 2016 sofre flutuações bastante significativas. Daí, e querendo equilibrar o estudo das simulações a realizar, estabelece-se que as previsões serão realizadas para última parte do ano, habitualmente para a secção equivalente aos últimos 25%.

Perceba-se que, existem diversos fatores que interferem na decisão do preço final. Como se trata de um mercado ibérico, não são apenas fatores nacionais que têm peso na atribuição do preço final da eletricidade. Para além disso, existem muitos fatores externos que podem ser influentes e que, nestas previsões sejam difíceis de pesar.

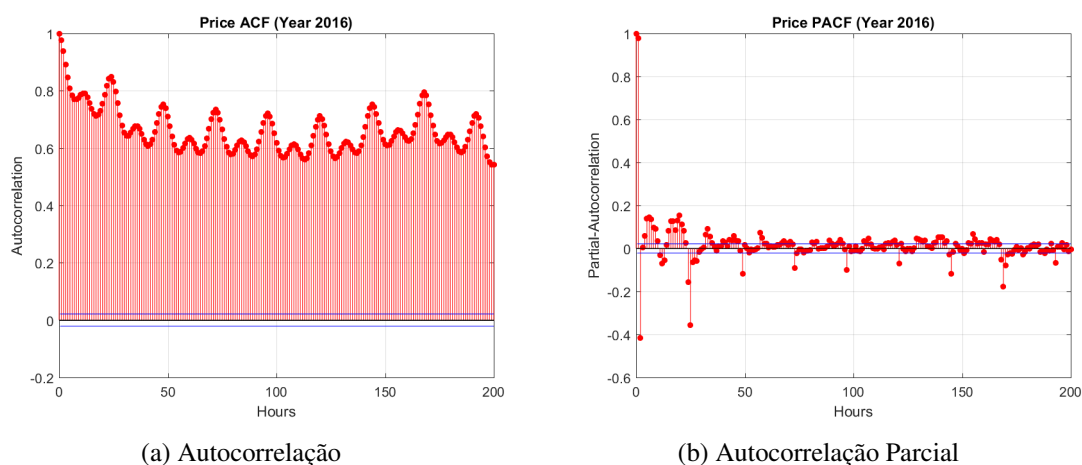


Figura 5.2: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2016.

A partir dos resultados demonstrados na Figura 5.2, facilmente se percebe a importância dos valores 24h, 48h e 168h anteriores. Ainda assim, como referido e como modelo base, ter-se-à em consideração apenas as variáveis do preço deslocadas 48h e 168h antes, garantindo que se impõe um horizonte que permite a execução de simulações que se enquadrem num ambiente próximo ao real.

Perante isto, e realizado o estudo desta série de preços da eletricidade de 2016 pelo método *Box-Jenkins*, como descrito em 2.4, resultam os seguintes valores de erro:

$$MAE_{Box-Jenkins_{2016}} = 6.774950259 \simeq 6.77 \text{ €} \quad (5.1)$$

Depois de entender que, com as variáveis de entrada deslocadas por 48 e 168 horas, o erro obtido por regressão linear (i.e., *Box-Jenkins*), é bastante razoável, pelo que a atribuição dos modelos para *GRNN* e *ANN* partirão deste, como base.

Assim, pretende-se realizar uma análise da evolução da curva do preço da eletricidade, assim como das ACF e PACF para os restantes anos, de forma a permitir que o modelo de base acima mencionado, seja válido e exequível para todos as séries em estudo.

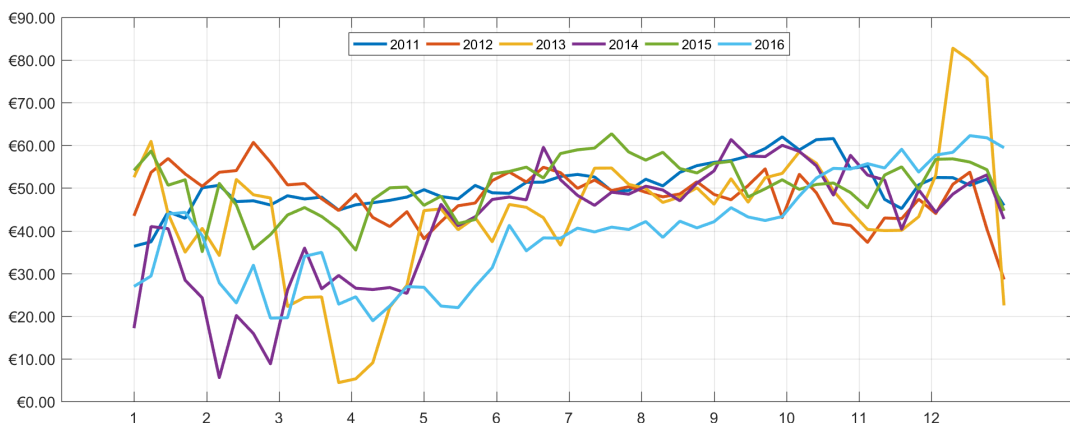


Figura 5.3: Evolução do preço médio semanal da eletricidade nos anos de 2011 a 2016 (€/MWh)

Como é possível verificar pela Figura 5.3, de uma forma aproximada (valores médios semanais), as curvas do preço da eletricidade nos diferentes anos tende a cometer flutuações razoáveis no primeiro trimestre e nos últimos meses do ano.

De forma a garantir que o padrão das variáveis a usar como modelo base será satisfatório, surge a necessidade de avaliar individualmente, as curvas do preço e as funções ACF e PACF, por ano.

Ora, relativamente aos outros anos em estudo, verifica-se sensivelmente o mesmo comportamento.

5.1.1.2 Preço da eletricidade em 2011

Esta secção sintetiza as análises efetuadas aos valores do preço da eletricidade em 2011, para balanço geral e identificação dos melhores atrasos a utilizar.

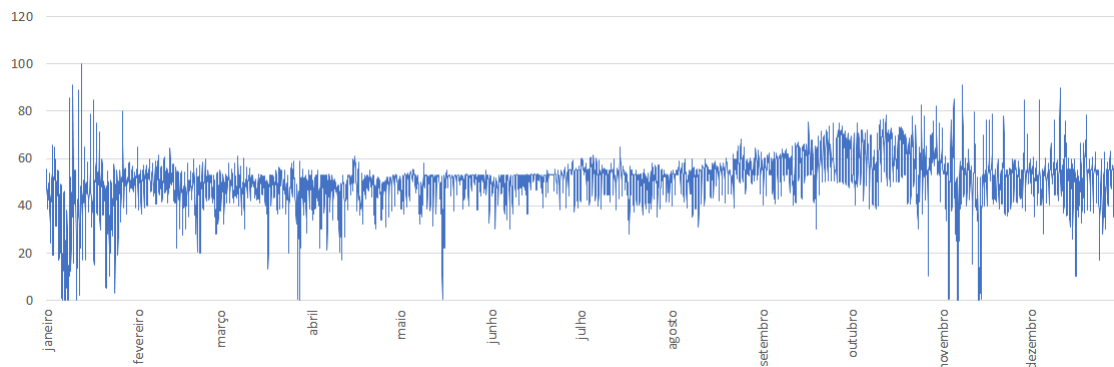


Figura 5.4: Evolução do preço da eletricidade em 2011 (€/MWh).

Como é possível verificar pela Figura 5.4, no ano de 2011, ainda que com alguns picos, verifica-se que o preço se mantém minimamente estável dentro de uma certa gama de valores. Ainda assim, sensivelmente a partir de setembro, nota-se que começa a surgir alguma flutuação, com picos elevados e outros até valores nulos.

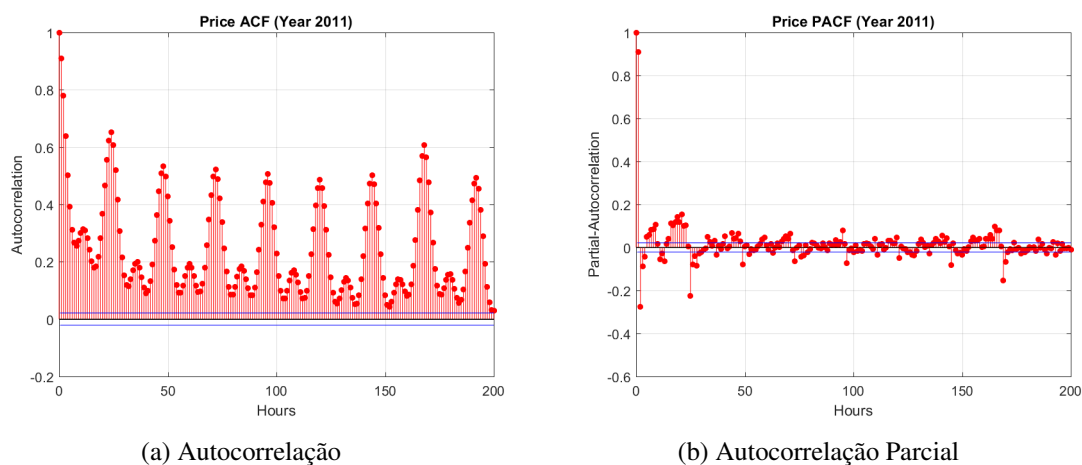


Figura 5.5: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2011.

Pela observação da Figura 5.5 facilmente se verifica uma forte correlação com os valores esperados (i.e., 24, 48 e 168 horas de atraso).

5.1.1.3 Preço da eletricidade em 2012

Esta secção sintetiza as análises efetuadas aos valores do preço da eletricidade em 2012, para balanço geral e identificação dos melhores atrasos a utilizar.

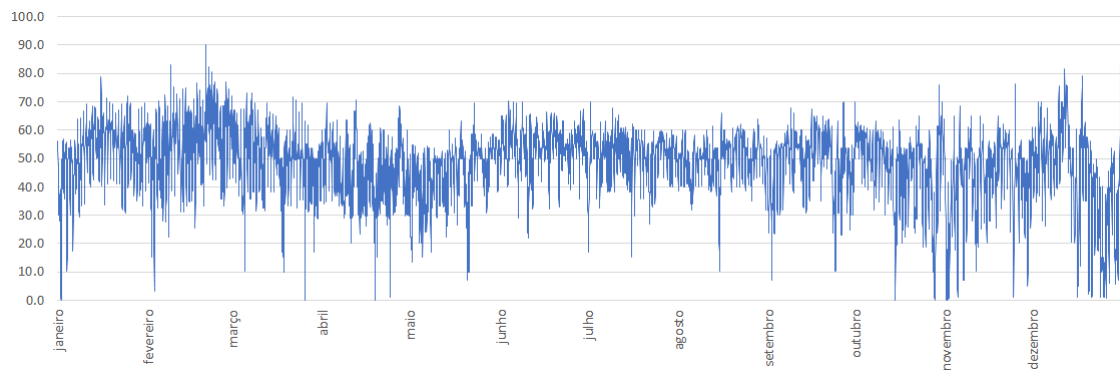


Figura 5.6: Evolução do preço da eletricidade em 2012 (€/MWh).

Verificando a Figura 5.6, percebe-se que 2012 se apresenta como um ano com maior flutuação de preços, com a presença de bastantes picos fora do normal, principalmente, e mais uma vez, no último trimestre, com oscilações dos 0 €/MWh até valores na casa dos 80 €/MWh.

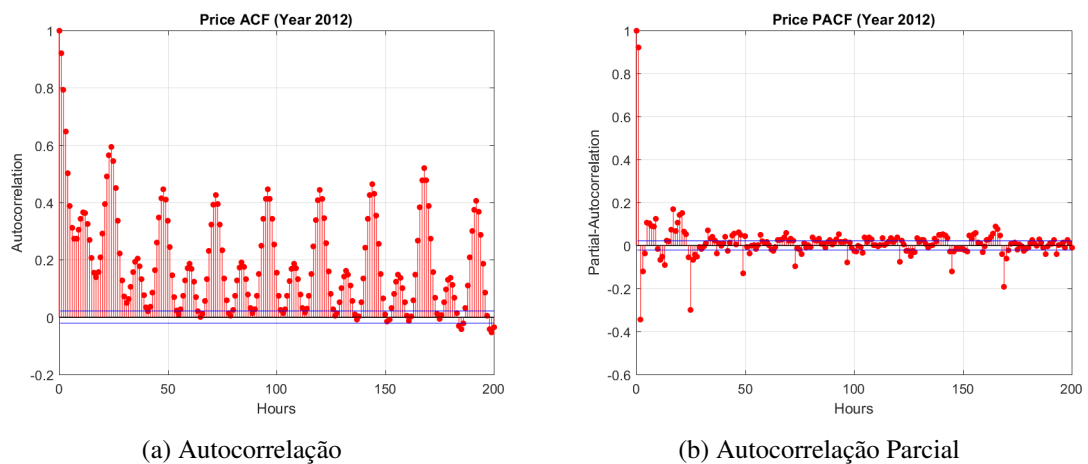


Figura 5.7: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2012.

Ainda que represente um ano mais volátil do que o anterior, pela análise das ACF e PACF apresentadas na Figura 5.7, também aparenta ter uma forte ligação com os valores associados a atrasos de 24, 48 e 168 horas anteriores.

5.1.1.4 Preço da eletricidade em 2013

Esta secção sintetiza as análises efetuadas aos valores do preço da eletricidade em 2013, para balanço geral e identificação dos melhores atrasos a utilizar.

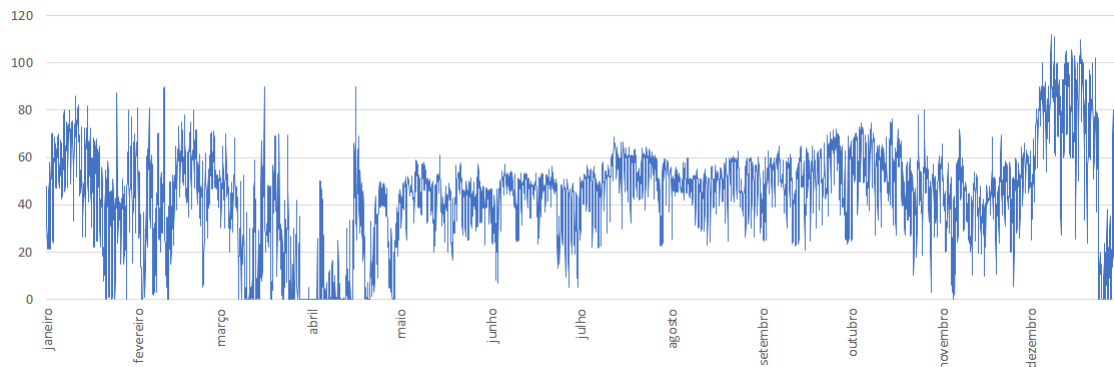


Figura 5.8: Evolução do preço da eletricidade em 2013 (€/MWh).

Ainda que já verificados anos com flutuações acentuadas, pela análise da Figura 5.8, percebe-se que o ano de 2013, relativamente à evolução do preço da eletricidade, apresenta oscilações mais proeminentes, com destaque claro para o primeiro e último trimestres, com o preço a atingir por diversas vezes valores nulos.

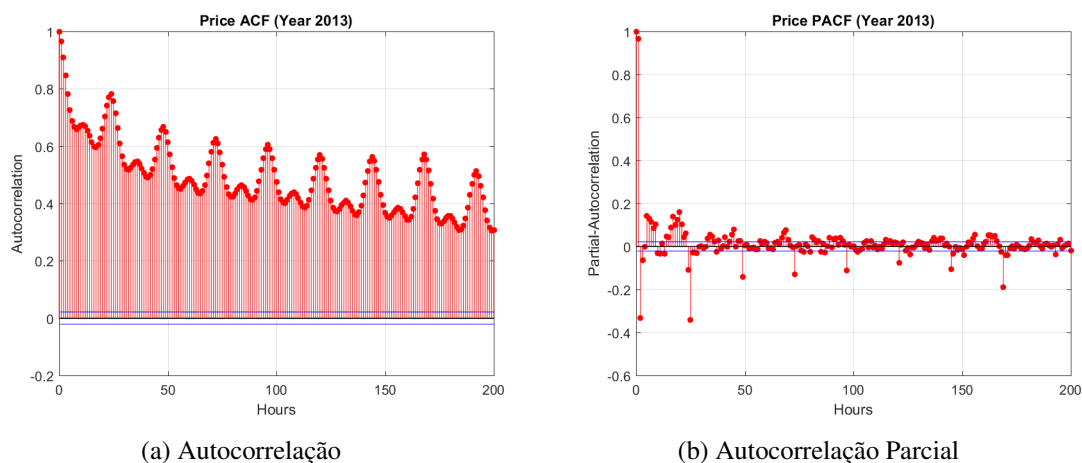


Figura 5.9: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2013.

Ainda assim, pela análise da Figura 5.9 se entende que se mantém uma fiel relação dos atrasos comuns, ou seja 24, 48 e 168 horas.

5.1.1.5 Preço da eletricidade em 2014

Esta secção sintetiza as análises efetuadas aos valores do preço da eletricidade em 2014, para balanço geral e identificação dos melhores atrasos a utilizar.

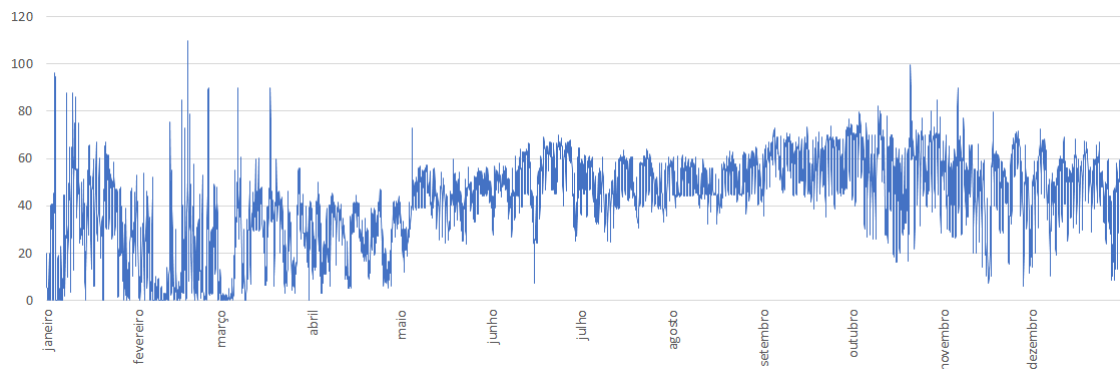


Figura 5.10: Evolução do preço da eletricidade em 2014 (€/MWh).

Dando continuidade ao oscilante final de 2013, o primeiro trimestre de 2014 também é caracterizado por acentuadas e repentinas mudança do preço da eletricidade. Ainda assim, ainda que com alguma flutuação, os meses seguintes apresentam uma evolução mais ou menos normal, com a exceção de alguns picos.

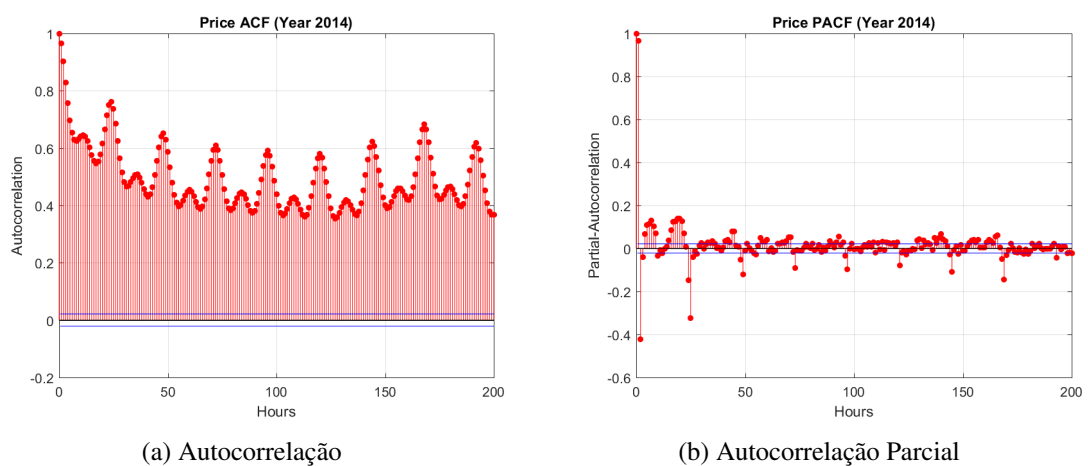


Figura 5.11: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2014.

Mantendo aquela que tem sido uma análise constante, também no ano de 2014, como é possível ver pela Figura 5.11 os atrasos de um dia, dois dias e uma semana, têm razoável correlação com o preço, pelo que também devem ser consideradas.

5.1.1.6 Preço da eletricidade em 2015

Esta secção sintetiza as análises efetuadas aos valores do preço da eletricidade em 2015, para balanço geral e identificação dos melhores atrasos a utilizar.

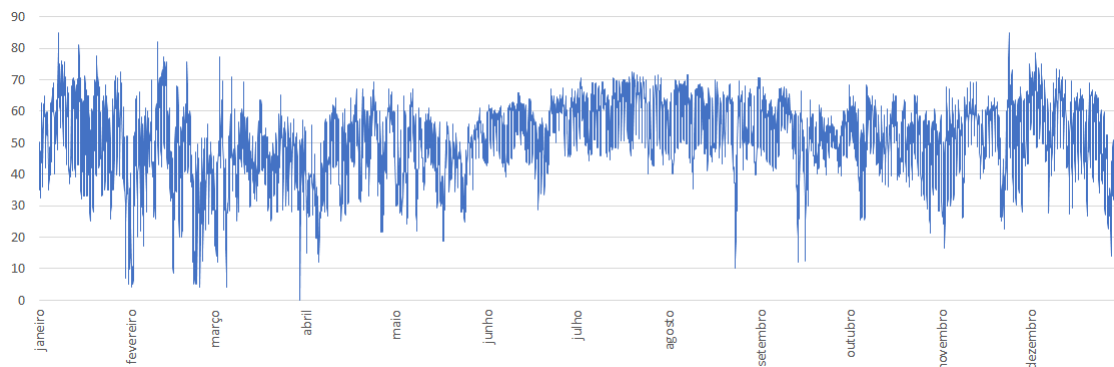


Figura 5.12: Evolução do preço da eletricidade em 2015 (€/MWh).

O ano de 2015, verificando a Figura 5.12 ainda que com algumas evidências de flutuações no primeiro trimestre, e com presença de alguns picos, garante uma evolução razoavelmente estável ao longo dos meses restantes.

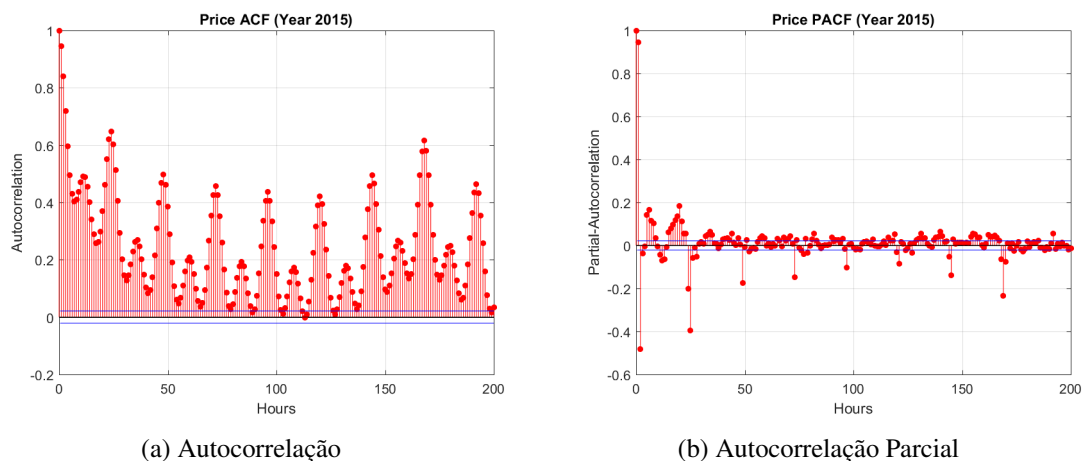


Figura 5.13: Funções de autocorrelação e autocorrelação parcial para o preço da eletricidade em 2014.

Finalmente, associando a análise da Figura 5.13 ao restantes anos, os atrasos com uma maior correlação, são mais uma vez os esperados, i.e., atrasos de 24, 48 e 168 horas.

5.1.1.7 Resumo da análise e resultados da previsão por *Box-Jenkins*

Em suma, analisados todos os anos de uma forma mais particular entende-se que, numa arena de variáveis a selecionar, aquelas que foram usadas para exemplificar o ano de 2016, reivindicam uma robustez suficiente para serem utilizadas nos restantes anos. Perceba-se que o *PACF* mostra que as variáveis mais informativas são as mais próximas do ponto a prever mas, como se pretende fazer a previsão para o dia seguinte, apenas são aceitáveis atrasos iguais ou superiores a 48 horas.

Ora, de forma a administrar uma referência nos valores dos erros, relativamente à previsão do preço da eletricidade, e tendo em conta mais uma vez o método estudado em 2.4, o modelo base a utilizar e respetivos erros (i.e., indicador de desempenho, no caso, *MAE*) são:

Modelo Base (Preço)

- Preço (h-48)
- Preço (h-168)

Estabelecido o modelo base para a previsão do preço, importa resolver e sintetizar os resultados da previsão do preço da eletricidade através do modelo *Box-Jenkins*, desta feita, utilizando o modelo base, e uma seleção de dados ordenada, com os primeiros 75% dos exemplos para cálculo dos pesos e os restantes 25% para verificação de resultados. Desta forma:

Tabela 5.1: Resultados de previsão do preço da eletricidade de 2011 a 2016, via *Box-Jenkins* (resultados por *MWh*).

Box-Jenkins	2011	2012	2013	2014	2015	2016
Modelo	Base (Preço)					
Amostragem	Ordenada (75% / 25%)					
MAE	6.27 €	9.67 €	12.89 €	8.95 €	6.55 €	6.77 €

Destaque-se na Tabela 5.1, a alguma variedade na gama de valores obtidos para o *MAE*, resultante talvez, da maior presença de fortes flutuações e também da presença de valores nulos. Perceba-se que a estratégia utilizada (regressão linear baseada nos valores passados da série) atribui sempre um peso independente das variáveis e, como tal, caso o preço a prever seja realmente nulo, a previsão torna-se mais pálida e dificilmente atingirá valores muito próximos de zero.

Ainda seja necessário desvendar o desempenho dos dois tipos de redes neuronais, é sensato notar que, ainda que normalmente mais ineficazes, as previsões via *Box-Jenkins* podem triunfar neste tipo de modelos simplistas, até porque a previsão via redes neuronais pode ser mais volátil, com poucas variáveis de entrada.

5.1.1.8 Simulações para identificação do melhor modelo

Perante isto, prende-se a necessidade de ir adicionando variáveis de entrada de forma a conseguir obter, por *GRNN* e *ANN*, os melhores resultados possíveis. Note-se que, ainda que seja possível e provável que a previsão via redes neuronais com o modelo base apresente resultados razoáveis, estas previsões estão associadas a uma fração de inconsistência. Isto é, ainda que numa previsão apresente bons resultados, é muito provável que as seguintes simulações rematem a previsão com erros elevados. Daí, e querendo claro realizar um correto exercício de previsão, atribui-se um maior peso à robustez do modelo.

Relativamente ao preço da eletricidade, importa claro ter em consideração as variáveis cíclicas (i.e., Mês, Dia da Semana - DDS - e Hora), os Feriados, a produção de energias renováveis e não-renováveis e também a previsão do consumo. Dessa forma, os modelos a serem testados terão como origem o modelo base, ao qual serão acrescentadas as variáveis referidas. Ainda mais, e de forma a querer sustentar uma significância de resultados ao longo de toda a série de dados, estabelece-se que a escolha do melhor modelo fará uso das técnicas de amostragem de dados de semanas misturadas e aleatória (ver Secção 4.3.2). Assim, facilmente se deteta a relação das variáveis adicionadas com uma amostra de dados global e não específica.

Tabela 5.2: Listagem de modelos a estudar para a previsão do preço da eletricidade

Modelo	Base	A	B	C
Mês		✓	✓	✓
Dia				
DDS		✓	✓	✓
Feriado			✓	✓
Feriado ($h - 48$)			✓	✓
Feriado ($h - 168$)			✓	✓
Hora		✓	✓	✓
Consumo ($h - 48$)				
Consumo ($h - 168$)				
Previsão do Consumo				✓
Preço ($h - 48$)	✓	✓	✓	✓
Preço($h - 168$)	✓	✓	✓	✓
Renováveis ($h - 48$)				✓
Renováveis ($h - 168$)				✓
Não-Renováveis ($h - 48$)				✓
Não-Renováveis ($h - 168$)				✓

Para a comparação entre tipos de redes neuronais e também entre modelos, referem-se os tempos de previsão (por tipo de rede) e a variação do *MAE* referente à previsão do preço da eletricidade, tendo como referência o modelo base. Toma-se como ponto de partida o ano de 2016.

5.1.1.9 Simulações para escolha do melhor modelo para 2016

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2016.

Tabela 5.3: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2016.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	5.88 €	5.81 €	5.45 €	5.55 €	6.06 €	4.94 €	3.10 €	3.63 €
MAPE'	14.35 %	14.17 %	13.78 %	14.03 %	14.78 %	12.04 %	7.80 %	9.13 %
RMSE	7.49 €	7.36 €	7.27 €	7.33 €	8.00 €	6.55 €	4.57 €	4.76 €
Tempo	2.10s	2.29s	2.03s	2.31s	3.56s	38.66s	3.85s	20.2s
Variação	-	-	-	-	↑ 0.18 €	↓ 0.87 €	↓ 2.35 €	↓ 1.92 €

Tabela 5.4: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2016.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	6.10 €	5.11	2.79 €	3.21 €	6.03 €	4.96 €	1.97 €	2.68 €
MAPE'	14.87 %	12.47 %	7.08 %	8.13 %	14.70 %	12.09 %	4.96 %	6.72 %
RMSE	8.08 €	6.70 €	4.20 €	4.27 €	8.04 €	6.43 €	3.13 €	3.62 €
Tempo	4.12s	1m 09s	4.35s	50.99s	7.19s	4m 01s	7.08s	2m 45s
Variação	↑ 0.22 €	↓ 0.70 €	↓ 2.66 €	↓ 2.34 €	↑ 0.15 €	↓ 0.85 €	↓ 3.48 €	↓ 2.87 €

Analisando as Tabelas 5.3 e 5.4 verifica-se então que o modelo C é aquele onde se verifica um melhor desempenho, pelo que será o modelo a utilizar para previsões do preço da eletricidade, neste ano.

Importa referir a importância da presença das variáveis cíclicas (modelo A) e a introdução dos feriados (modelo B). Finalmente, de forma a reduzir o erro, introduz-se no modelo a previsão de consumo e a produção de energias renováveis e não renováveis que, ao que tudo indica, melhoram os modelos anteriores.

5.1.1.10 Simulações para escolha do melhor modelo para 2015

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2015.

Tabela 5.5: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2015.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	5.87€	5.70€	6.78€	6.81€	6.95€	5.94€	4.08€	4.99€
MAPE'	11.28 %	10.95 %	13.55 %	13.61 %	13.36 %	11.43 %	8.15 %	9.97 %
RMSE	7.77€	7.45€	8.98€	9.01€	9.42€	8.19€	5.97€	6.58€
Tempo	2.24s	3.09s	2.73s	2.36s	3.77s	25.24s	4.8s	19.23s
Variação	-	-	-	-	↑ 1.08€	↑ 0.24€	↓ 2.70€	↓ 1.82€

Tabela 5.6: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2015.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	6.92€	6.18€	3.78€	4.43€	6.05€	5.75€	2.79€	4.03€
MAPE'	13.30 %	11.87 %	7.50 %	8.80 %	11.64 %	11.06 %	5.57 %	8.05 %
RMSE	9.38€	8.38€	5.78€	5.87€	8.37€	7.50€	4.36€	5.42€
Tempo	4.72s	1m 04s	4.8s	5.87s	6.54s	2m 24s	7.54s	1m 43s
Variação	↑ 1.05€	↑ 0.48€	↓ 3.00€	↓ 2.38€	↑ 0.18€	↑ 0.05€	↓ 3.99€	↓ 2.78€

Como é possível verificar pelas Tabelas 5.5 e 5.6, percebe-se uma pequena perda de desempenho nos modelos A e B, com o modelo C, ainda que também associado a leves perdas, a ser aquele que se destaca. Entenda-se que a redução de desempenho na seleção semanal neste último modelo é mínima, com a seleção aleatória a destacar-se pela positiva, pelo que se escolhe este como o modelo utilizar.

5.1.1.11 Simulações para escolha do melhor modelo para 2014

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2015.

Tabela 5.7: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2014.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	9.24 €	9.24 €	8.59 €	8.78 €	10.08 €	8.50 €	4.93 €	6.26 €
MAPE'	21.79 %	21.81 %	20.24 %	20.70 %	23.78 %	20.06 %	11.63 %	14.78 %
RMSE	12.14 €	12.13 €	11.37 €	11.49 €	14.19 €	11.33 €	7.60 €	8.31 €
Tempo	2.03s	2.22s	2.19s	2.52s	3.95s	25.08s	3.77s	21.45s
Variação	-	-	-	-	↑ 0.84 €	↓ 0.74 €	↓ 3.66 €	↓ 2.52 €

Tabela 5.8: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2014.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	10.20 €	9.02 €	4.58 €	5.69 €	10.39 €	9.51 €	3.12 €	5.26 €
MAPE'	24.07 %	21.29 %	10.80 %	13.42 %	24.53 %	22.43 %	7.35 %	12.40 %
RMSE	14.34 €	12.00 €	7.20 €	7.66 €	14.66 €	12.44 €	5.51 €	7.13 €
Tempo	4.2s	1m 23s	4.74s	55.04s	6.57s	2m 43s	6.52s	3m 37s
Variação	↑ 0.96 €	↓ 0.22 €	↓ 4.01 €	↓ 3.09 €	↑ 1.15 €	↑ 0.27 €	↓ 5.47 €	↓ 3.52 €

Mais uma vez, vendo pelas Tabelas 5.7 e 5.8 também no ano de 2014 o modelo a utilizar será o modelo C. Nota-se, de facto, uma leve perda de desempenho com a adição de variáveis no modelo semanal. Porém, as diferenças são leves ao ponto de ser razoável considerar as enormes melhorias causadas na seleção aleatória, notoriamente visíveis no modelo C.

5.1.1.12 Simulações para escolha do melhor modelo para 2013

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2013.

Tabela 5.9: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2013.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	11.18 €	10.83 €	10.48 €	10.70 €	12.94 €	11.19 €	5.74 €	7.30 €
MAPE'	25.74 %	24.94 %	24.18 %	24.69 %	29.80 %	25.76 %	13.28 %	16.89 %
RMSE	15.45 €	14.96 €	13.95 €	14.15 €	18.90 €	15.89 €	9.44 €	9.99 €
Tempo	1.69s	3.04s	2.56s	2.27s	3.77s	24.88s	3.72s	21.86s
Variação	-	-	-	-	↑ 1.76 €	↑ 0.36 €	↓ 4.74 €	↓ 3.40 €

Tabela 5.10: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2013.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	13.36 €	12.00 €	5.29 €	7.10 €	12.89 €	11.58 €	3.42 €	6.16 €
MAPE'	30.76 %	27.63 %	12.13 %	16.28 %	29.69 %	26.66 %	7.76 %	13.99 %
RMSE	19.23 €	17.28 €	9.10 €	10.05 €	18.32 €	15.99 €	5.88 €	8.61 €
Tempo	4.35s	1m 23s	4.53s	47.84s	6.13s	4m 05s	6.61s	2m 25s
Variação	↑ 2.18 €	↑ 1.17 €	↓ 5.19 €	↓ 3.60 €	↑ 1.71 €	↑ 0.75 €	↓ 7.06 €	↓ 4.54 €

Como verificado acima, o ano de 2013 comporta-se, de facto, de uma forma inconsistente. Ainda assim, salvo a perda de desempenho nas transições para o modelo A e B na seleção semanal (ver Tabelas 5.9 e 5.10), as melhorias impostas pelas transições de modelos de uma forma global são notórias e, como tal, também se fará uso do modelo C, onde as perdas na seleção semanal, associadas aos ganhos da seleção aleatória tornam, de uma forma global, o modelo C como o mais vantajoso.

5.1.1.13 Simulações para escolha do melhor modelo para 2012

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2012.

Tabela 5.11: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2012.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	6.86 €	6.79 €	7.00 €	7.02 €	8.08 €	7.10 €	4.60 €	5.28 €
MAPE'	14.46 %	14.32 %	14.54 %	14.60 %	17.05 %	14.96 %	9.55 %	10.90 %
RMSE	9.67 €	9.57 €	9.70 €	9.74 €	11.24 €	10.24 €	6.77 €	7.28 €
Tempo	2.7s	1.71s	3.94s	2.01s	3.96s	14.5s	4.44s	13.45s
Variação	-	-	-	-	↑ 1.22 €	↑ 0.31 €	↓ 2.40 €	↓ 1.74 €

Tabela 5.12: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2012.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	8.07 €	6.52 €	4.15 €	4.68 €	7.11 €	6.28 €	3.06 €	4.27 €
MAPE'	17.02 %	13.75 %	8.58 %	9.69 %	14.99 %	13.24 %	6.39 %	8.91 %
RMSE	11.05 €	9.09 €	6.11 €	6.41 €	10.19 €	8.68 €	4.68 €	5.93 €
Tempo	4.5s	1m 08s	5.4s	33.35s	7.6s	1m 18s	7.43s	1.20s
Variação	↑ 1.21 €	↓ 0.27 €	↓ 2.85 €	↓ 2.34 €	↑ 0.25 €	↓ 0.51 €	↓ 3.94 €	↓ 2.75 €

O ano de 2012 apresenta uma evolução de desempenho bastante razoável. Excetuando as notáveis descidas de performance nos modelos A e B, para a seleção semanal, as transições de modelo provocam, de uma forma geral, melhorias no desempenho de uma forma gradual e constante (ver Tabelas 5.11 e 5.12). Como tal, opta-se claramente pela utilização do modelo C.

5.1.1.14 Simulações para escolha do melhor modelo para 2011

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do preço da eletricidade para o ano de 2011.

Tabela 5.13: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos Base e A, relativos ao ano de 2011.

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	4.37 €	4.31 €	4.41 €	4.43 €	4.79 €	4.41 €	3.33 €	3.62 €
MAPE'	8.48 %	8.36 %	8.66 %	8.70 %	9.29 %	8.56 %	6.55 %	7.10 %
RMSE	6.93 €	6.75 €	6.83 €	6.83 €	7.58 €	7.01 €	5.24 €	5.34 €
Tempo	3.2s	1.91s	3.66s	1.6s	4.24s	15.31s	4.59s	9.28
Variação	-	-	-	-	↑ 0.42 €	↑ 0.10 €	↓ 1.08 €	↓ 0.81 €

Tabela 5.14: Resultados dos testes de previsão do preço da eletricidade, por redes neuronais, para os modelos B e C, relativos ao ano de 2011.

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAE	5.26 €	4.64 €	3.23 €	3.54 €	5.11 €	4.67 €	2.27 €	3.35 €
MAPE'	10.20 %	9.01 %	6.40 %	7.01 %	9.91 %	9.05 %	4.46 %	6.59 %
RMSE	8.20 €	7.18 €	5.10 €	5.35 €	8.10 €	6.94 €	3.90 €	5.02 €
Tempo	4.4s	39.85s	5.5s	27.04s	6.98s	1m 52s	7.31s	1m 27s
Variação	↑ 0.89 €	↑ 0.33 €	↓ 1.18 €	↓ 0.89 €	↑ 0.74 €	↑ 0.36 €	↓ 2.14 €	↓ 1.08 €

Na situação do ano de 2011, como se consegue perceber pela análise das Tabelas 5.13 e 5.14, ainda que com sobressaltos mínimos na evolução do modelo, compreende-se que, de uma forma geral, a adição daquelas variáveis coordena melhorias no desempenho, pelo que também se utilizará o modelo C.

5.1.1.15 Resumo das simulações

Analisados a fundo os anos em estudo, e também de forma a uniformizar o trabalho a executar, opta-se por impor o modelo C como modelo de previsão do preço da eletricidade para os anos de 2011 a 2016. Clarifique-se que, ainda que nalguns modelos, para alguns anos, se tenha verificado um razoável desempenho, os resultados destas previsões são, de certa forma, demasiado voláteis para o tipo de previsão a executar. Julga-se mais correta a execução deste exercício quando se garante que o modelo a utilizar responde sempre de uma forma equilibrada e constante a um mesmo conjunto de dados.

Tabela 5.15: Resumo dos modelos a utilizar para a previsão do preço da eletricidade, por ano.

Ano	2011	2012	2013	2014	2015	2016
Modelo	C	C	C	C	C	C

5.1.2 Consumo de Energia Elétrica

Assim como planeado na escolha do melhor modelo para a previsão do preço da eletricidade, a escolha do melhor modelo de previsão do consumo de energia elétrica, também deve seguir um rumo idêntico. Assim, como primeira etapa, analisam-se as curvas do consumo de energia ao longo do ano, para cada ano, e daí se verificam quais devem ser aquelas variáveis que constituem o modelo base de previsão. De seguida, são adicionadas aquelas variáveis que podem ter relação com a grandeza a prever, até que se obtenha o modelo ideal.

Mais uma vez, querendo garantir que a previsão para o dia seguinte é realizada com a antecedência necessária, opta-se por fixar horizontes temporais com um atraso mínimo referentes a quarenta e oito horas (i.e., dois dias). Ora, impostos dados referentes, no pior dos casos, a dois dias anteriores, garante-se que se consiga fazer uma previsão para o dia seguinte com uma antecedência razoável.

5.1.2.1 Consumo de energia elétrica em 2016

Usando como exemplo o ano de 2016, realiza-se uma previsão simples por regressão linear e através do método de *Box-Jenkins* (ver Secção 2.4) utilizando como entradas os valores da variável a prever, mas obviamente deslocadas consoante os horizontes temporais a considerar. Ainda assim, de forma a garantir que os atrasos escolhidos estão em conformidade com a variável a prever, realiza-se também uma análise das funções ACF e PACF.

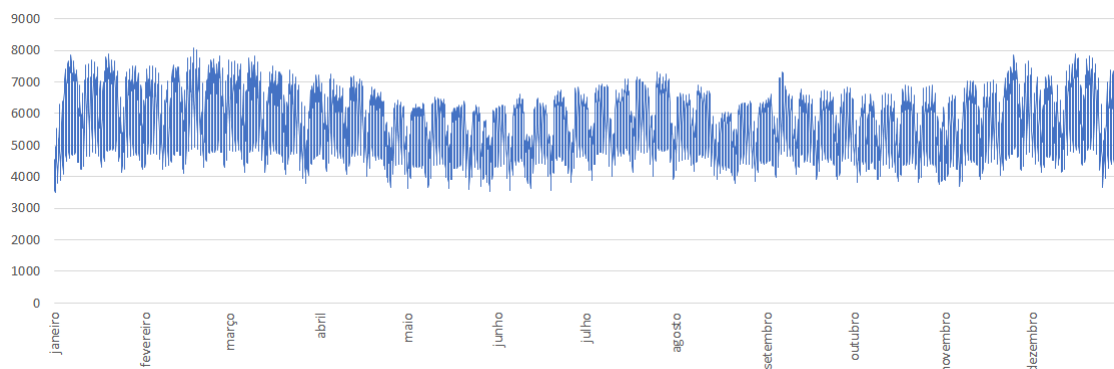


Figura 5.14: Evolução do consumo de energia elétrica em 2016 (*MWh*).

Verificando a Figura 5.14, e contrariamente ao que ocorria com o preço da eletricidade, percebe-se claramente que a evolução do consumo de energia elétrica é bastante mais equilibrada e tende a respeitar um padrão relativamente estável. Ou seja, ainda que sujeito a condições externas (e.g., estação do ano, dia da semana, etc.), os valores do consumo tendem a oscilar sempre dentro de uma gama minimamente constante de valores. Neste caso, verifica-se que o consumo de energia elétrica, ao longo do ano, se situa entre os 3500 a 8000 *MWh*.

Percebe-se facilmente que nas épocas de menor temperatura, o consumo tem tendência a aumentar um pouco, ocorrendo o contrário nas épocas de maior calor, embora já se comece a identificar uma pequena subida a meio do Verão, algo que não acontecia no passado. Para além disso, como o diagrama de carga é razoavelmente constante, é comum referir que este deve respeitar uma forma semelhante ao dos mesmos dias semanais.

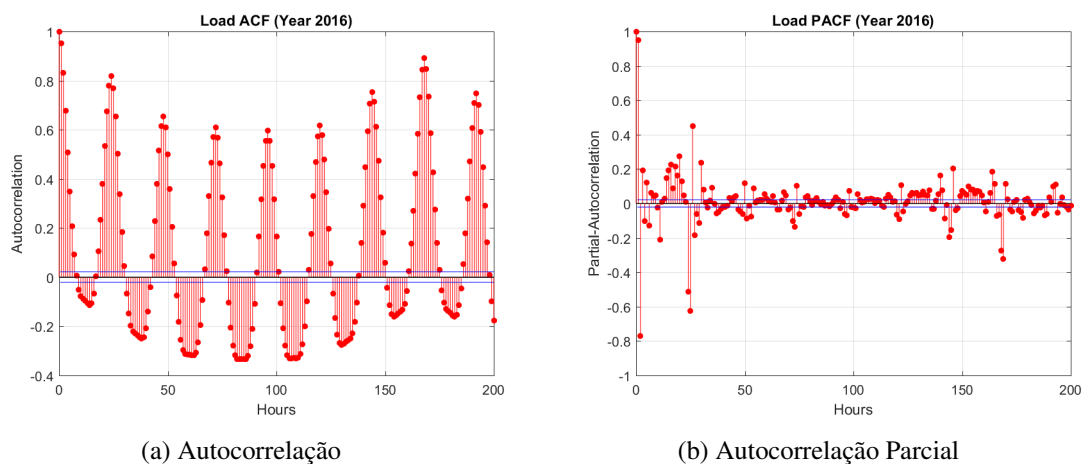


Figura 5.15: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2016.

Verificando a Figura 5.15, facilmente se percebe que os valores com um atraso de 24 em 24 horas têm elevada relevância destacando-se, ainda assim, os atrasos de 24 e 168 horas (i.e., um dia e uma semana antes). Como referido, o valor do consumo referente ao mesmo dia semanal

é crucial para uma boa previsão. Perceba-se que, ainda que o atraso de 24 horas seja importante para a previsão, querendo estabelecer um modelo que seja exequível para previsões reais, opta-se por escolher como base, os atrasos de 48 e 168 horas.

Assim, e de forma a tomar uma referência para previsão do consumo de energia elétrica, pelo método *Box-Jenkins*, e para o ano de 2016, resulta o seguinte valor de erro:

$$MAPE_{Box-Jenkins_{2016}} = 0.050650046 \simeq 5.065 \% \quad (5.2)$$

Verificado que, com os atrasos de 48 e 168 horas, os resultados obtidos por regressão linear (i.e., *Box-Jenkins*) são bons, impõe-se que o modelo base da previsão por *GRNN* e *ANN*, do consumo de energia elétrica, será este. Desta forma, e antes de se proceder à escolha do modelo final para previsão do consumo para cada ano, existe necessidade também de se analisar as curvas, funções ACF e PACF de cada um destes anos, de forma individual.

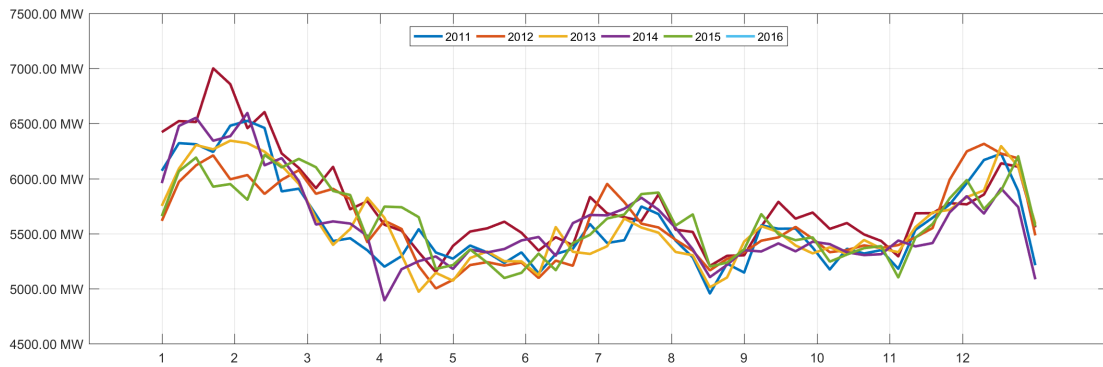


Figura 5.16: Evolução do consumo médio semanal de energia elétrica e nos anos de 2011 a 2016 (*MWh*)

Analisando a Figura 5.16, facilmente se percebe que a evolução do consumo ao longo de um ano tende claramente a respeitar um certo padrão. Ora, percebe-se que o período de menor temperatura (i.e., primeiros quatro e últimos dois meses) representa um aumento no consumo, acompanhado por uma ligeira subida na época de maior calor (i.e., entre junho e agosto). De uma forma geral, e salvo flutuações singulares, a evolução do consumo ao longo de um ano, nos períodos a analisar, parece ser relativamente equilibrada e constante.

O padrão de consumo tende a ser mais ou menos constante ao longo de diferentes dias. Por isto, e salvo circunstâncias especiais (e.g., dias festivos), a curva de consumo de dois dias, mais ou menos semelhantes, deve ser, ela também, semelhante. Ainda assim, surge a necessidade de analisar de forma individual (i.e., por ano), a coerência dos atrasos mencionados acima para os restantes anos.

5.1.2.2 Consumo de energia elétrica em 2011

Esta secção sintetiza as análises efetuadas aos valores do consumo de energia elétrica em 2011, para balanço geral e identificação dos melhores atrasos a utilizar.

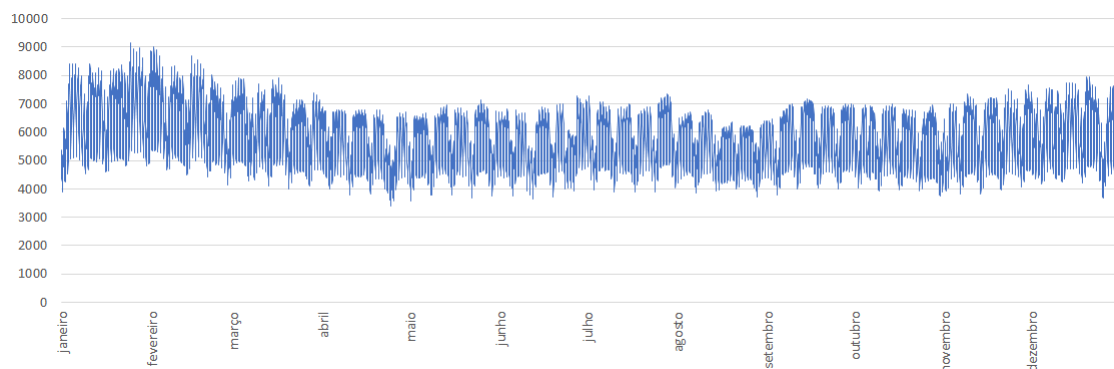


Figura 5.17: Evolução do consumo de energia elétrica em 2011 (*MWh*).

Como é possível verificar pela Figura 5.17, no ano de 2011, o consumo de energia elétrica respeita o padrão habitual, com uma ligeira subida no início do ano, mas sempre coerente com uma gama de valores, no caso, entre os 3500 e 9000 *MWh*, salvo exceção.

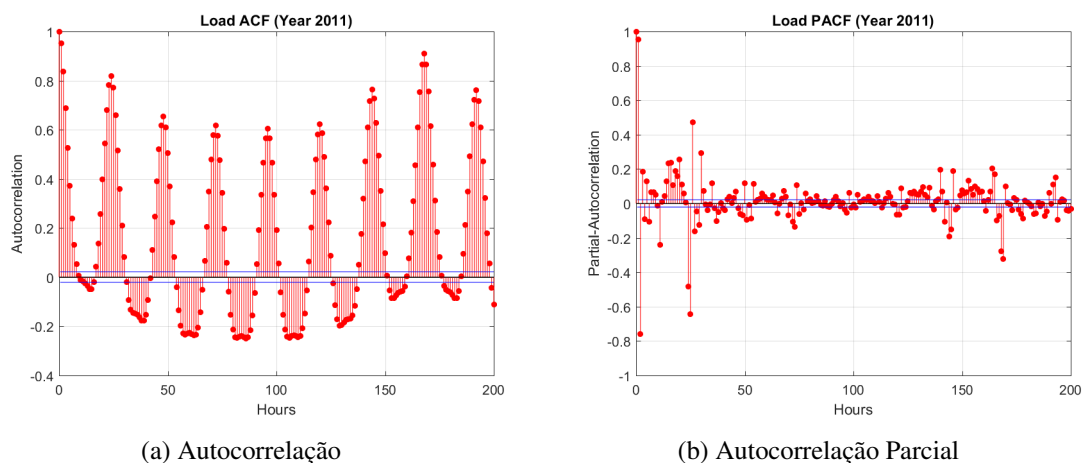


Figura 5.18: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2011.

Através da Figura 5.18 mais uma vez se percebe que os atrasos de 24 e 168 horas são aqueles que se destacam. Ainda assim, como estudado, não se considerará a deslocação das 24 horas, de forma a garantir que é respeitada a antecendência para previsões que é requerida. Desta forma considerar-se-à, no seu lugar, o atraso de 48 horas.

5.1.2.3 Consumo de energia elétrica em 2012

Esta secção sintetiza as análises efetuadas aos valores do consumo de energia elétrica em 2012, para balanço geral e identificação dos melhores atrasos a utilizar.

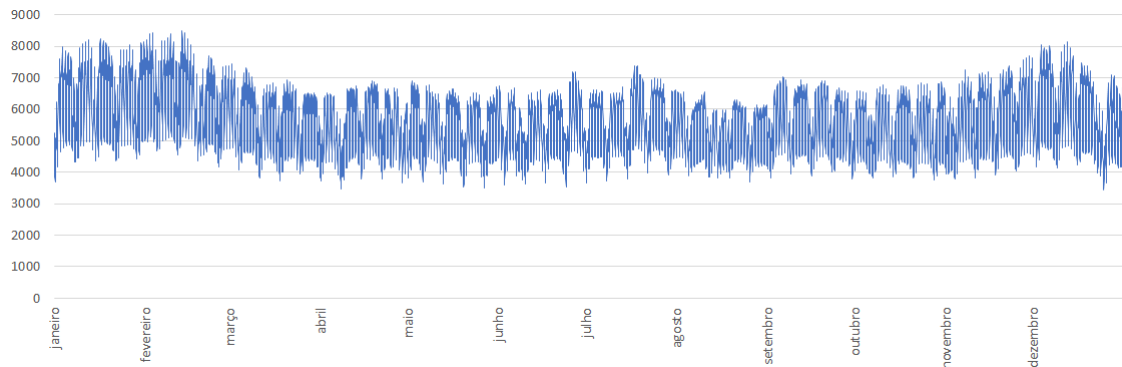


Figura 5.19: Evolução do consumo de energia elétrica em 2012 (MWh).

Mais uma vez, desta feita no ano de 2012, como se verifica pela Figura 5.19, existe um aumento de consumo no início e no final do ano, juntamente com pequenas subidas na época de maior calor, mais precisamente entre junho e agosto. O consumo apresenta-se sem oscilações anormais, pelo que é possível afirmar que respeita o padrão esperado.

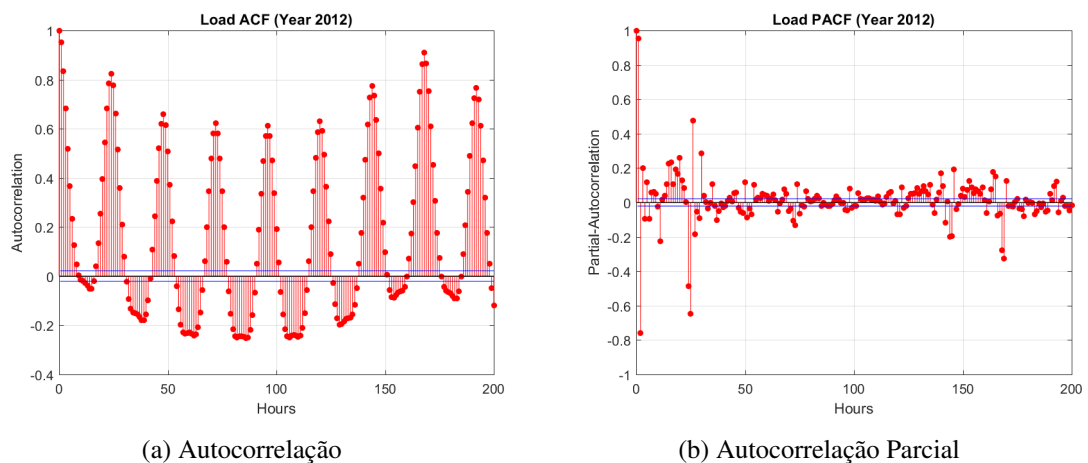


Figura 5.20: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2012.

Aqui também, analisando a Figura 5.20, se deteta importância nos valores associados a deslocamentos de 24 e 168, um dia e uma semana, respetivamente. Também como solucionado nos casos anteriores, considera-se o atraso de 48 horas ao invés do de 24 horas.

5.1.2.4 Consumo de energia elétrica em 2013

Esta secção sintetiza as análises efetuadas aos valores do consumo de energia elétrica em 2013, para balanço geral e identificação dos melhores atrasos a utilizar.

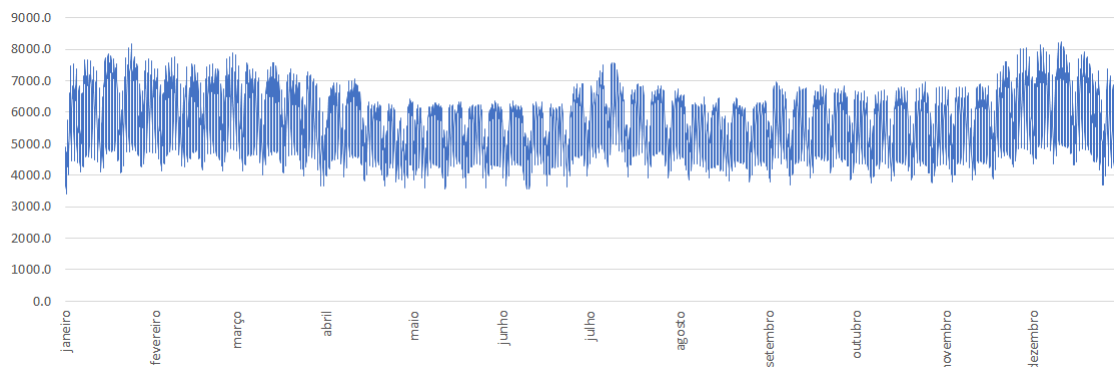


Figura 5.21: Evolução do consumo de energia elétrica em 2013 (MWh).

Ainda que de acordo com o esperado, o ano de 2013, como se verifica pela Figura 5.21, é razoavelmente constante e tratável até por blocos. Note-se que os períodos de maior frio acarretam uma subida de quase 2000 MWh para o valor de pico, quando comparados com períodos mais amenos, como aqueles verificados entre maio e junho. Mais uma vez, também no verão, ou mais precisamente no mês de julho, nota-se um aumento significativo, talvez associado ao aumento de temperatura, normal nessa época do ano.

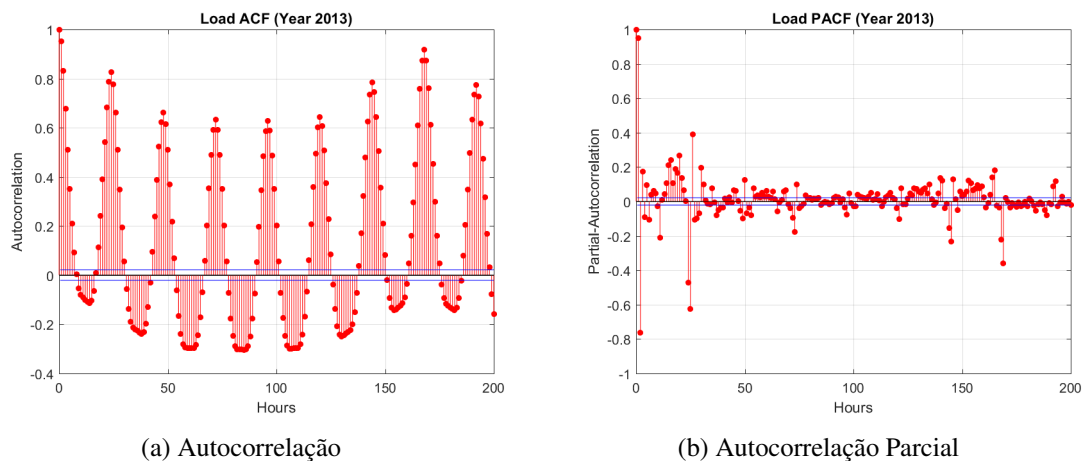


Figura 5.22: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2013.

Sem muitas diferenças relativamente ao já analisado, analisando a Figura 5.22, rapidamente se entende que os atrasos habituais (i.e., de 24 e 168 horas) são aqueles com maior correlação. Aqui também, graças à antecedência requerida, serão considerados os atrasos de 48 e 168 horas.

5.1.2.5 Consumo de energia elétrica em 2014

Esta secção sintetiza as análises efetuadas aos valores do consumo de energia elétrica em 2014, para balanço geral e identificação dos melhores atrasos a utilizar.

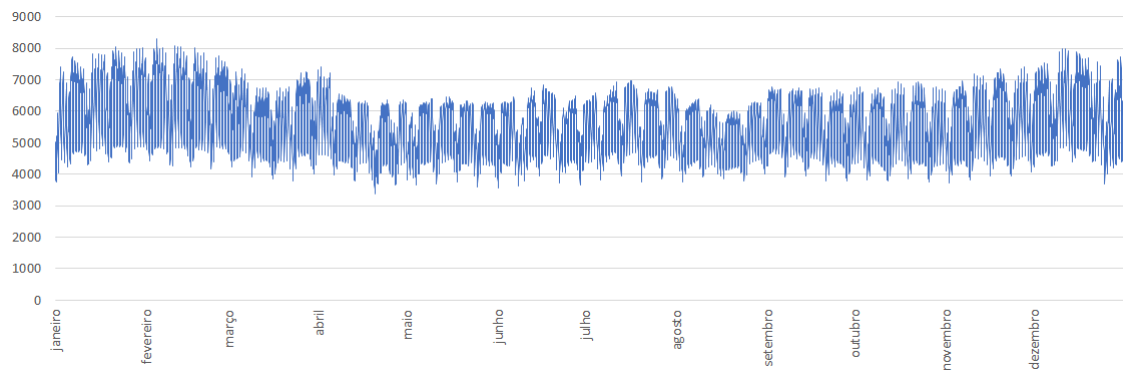


Figura 5.23: Evolução do consumo de energia elétrica em 2014 (MWh).

De forma relativamente idêntica ao que ocorreu nos anos já estudados, a evolução do consumo em 2014, mais uma vez, tente a ser estável e equilibrada, com aumentos nas épocas de maior frio e com uma leve subida entre julho e agosto (ver Figura 5.23).

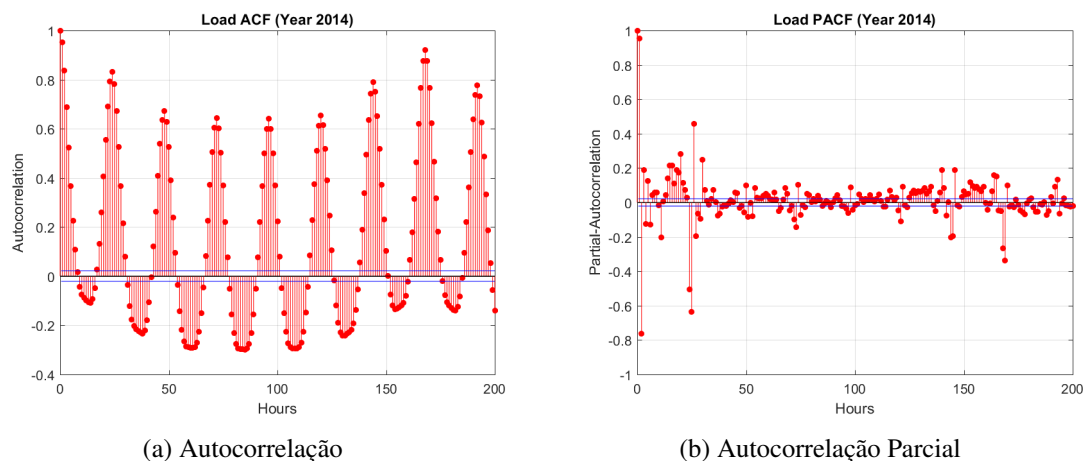


Figura 5.24: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2014.

Sem grandes expectativas de mudança, uma rápida análise da Figura 5.24 demonstra que os atrasos de 24 e 168 continuam a ser coerentes com os anos já analisados. Ora, mais uma vez privilegiando a antecedência da previsão, consideram-se os atrasos de 48 e 168 horas.

5.1.2.6 Consumo de energia elétrica em 2015

Esta secção sintetiza as análises efetuadas aos valores do consumo de energia elétrica em 2015, para balanço geral e identificação dos melhores atrasos a utilizar.

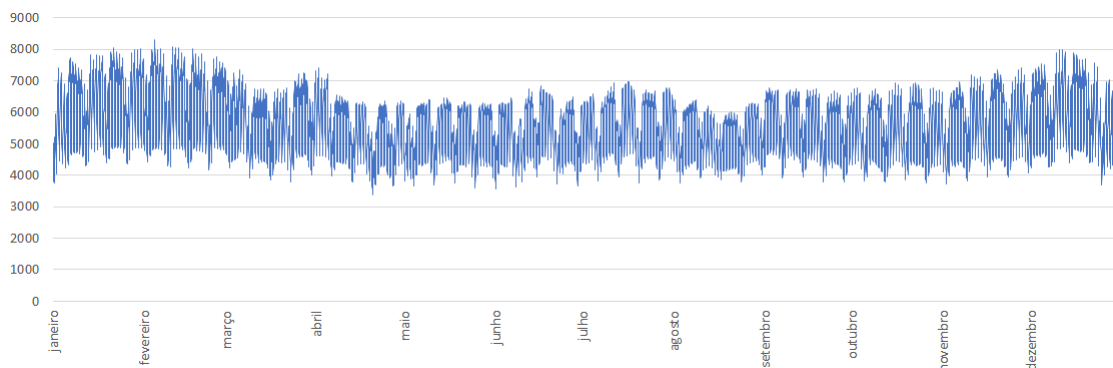


Figura 5.25: Evolução do consumo de energia elétrica em 2015 (MWh).

Finalmente, relativamente ao ano de 2015, analisando a Figura 5.25, percebe-se facilmente que o panorama de estabilidade e equilíbrio do consumo de energia elétrica se mantém. Sem fugir claramente aos limites habituais (i.e., situa-se sempre entre os 3000 e os 9000 MWh), respeita aquilo que tem sido verificado nos restantes anos.

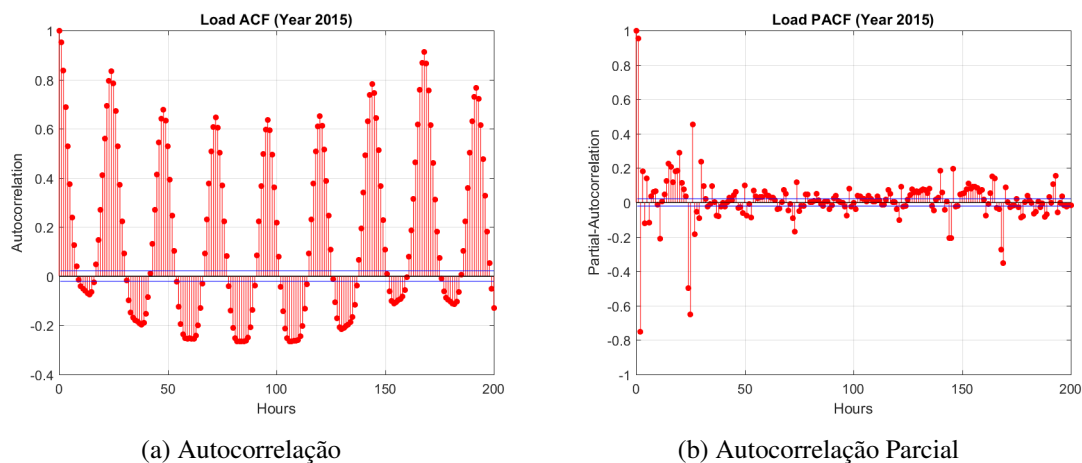


Figura 5.26: Funções de autocorrelação e autocorrelação parcial para o consumo de energia elétrica em 2015.

Também neste ano, por uma rápida análise da Figura 5.26, se deteta que os deslocamentos a aplicar para as previsões do consumo em 2015 serão também de 48 e 168 horas, tendo em consideração, claro, as limitações do atraso de 24 horas.

5.1.2.7 Resumo da análise e resultados da previsão por Box-Jenkins

De uma forma geral, a análise da evolução do consumo de energia elétrica ao longo dos anos de 2011 até 2016 demonstra que, salvo exceção, se trata de uma série bastante estável. Como tal, certificados que os atrasos são comuns a todos os anos, e querendo impor uma referência para a previsão via *GRNN* e *ANN*, faz-se uma previsão via *Box-Jenkins* (ver Secção 2.4), utilizando um indicador de desempenho (no caso, *MAPE*), e o modelo base para previsão do consumo, que contempla:

Modelo Base (Consumo)

- Consumo (h-48)
- Consumo (h-168)

Esclarecido qual o modelo base para a previsão do consumo, importa sintetizar os resultados da previsão realizados com recurso ao método de Box-Jenkins. Ora, a previsão é realizada utilizando o modelo base para o consumo, com uma seleção de dados ordenada com os primeiros 75% dos exemplos para cálculo dos pesos e os restantes 25% para verificação de resultados. Daqui:

Tabela 5.16: Resultados da previsão do consumo de energia elétrica de 2011 a 2016, via *Box-Jenkins*

Box-Jenkins	2011	2012	2013	2014	2015	2016
Modelo	Base (Consumo)					
Amostragem	Ordenada (75% / 25%)					
MAPE	4.020 %	4.646 %	3.579 %	3.826 %	3.758 %	5.065 %

Analisando os valores apresentados na Tabela 5.16 facilmente se percebe que, contrariamente ao que acontecia na previsão do preço da eletricidade, os resultados da previsão por Box-Jenkins, para além de serem bastante razoáveis, não oscilam muito entre eles. Note-se que, como o indicador utilizado foi o *MAPE*, uma variação unitária, não é muito significativa, até porque os resultados por si só, já são de bom nível.

Ainda que os algoritmos de amostragem utilizados não sejam utilizáveis em situações reais de uma forma direta, são bastante importantes na verificação do comportamento das séries. Ora, por exemplo a seleção por semanas misturadas é capaz de realizar um estudo da série em blocos ao longo de todo o ano, enquanto que uma seleção ordenada se fixa apenas nos valores finais de cada ano. Assim, ao invés de analisar apenas uma secção específica do ano, que por si só pode ser uma secção de mais difícil estimação, realiza uma previsão seccionada, e daí se consegue entender que as variáveis introduzidas seriam importantes para a previsão por redes neuronais.

5.1.2.8 Simulações para identificação do melhor modelo

Perceba-se ainda que, relativamente ao consumo de energia elétrica, e tratando-se de uma série razoavelmente estável, percebe-se que é um alvo cujas previsões se podem apresentar com melhores resultados gerais. Como referido, como a curva de consumo respeita de uma forma muito direta a hora do dia e o dia da semana, importa impor ao modelo base este tipo de entradas, assim como outras variáveis cíclicas (e.g. Mês). Importa também verificar o impacto dos feriados, até porque o consumo num feriado tende a ser distinto de um dia útil normal. Desta forma, e querendo obter o melhor modelo para previsão, os modelos a serem testados serão baseados no modelo base, ao qual serão acrescentadas e/ou retiradas, se necessários, novas variáveis.

Tabela 5.17: Listagem de modelos a estudar para a previsão do consumo de energia elétrica

Modelo	Base	A	B	C
Mês		✓	✓	✓
Dia				
DDS		✓	✓	✓
Feriado			✓	✓
Feriado ($h - 48$)			✓	✓
Feriado ($h - 168$)			✓	✓
Hora		✓	✓	✓
Consumo ($h - 48$)	✓	✓	✓	✓
Consumo ($h - 168$)	✓	✓	✓	✓
Previsão do Consumo				
Preço ($h - 48$)				
Preço($h - 168$)				
Renováveis ($h - 48$)				✓
Renováveis ($h - 168$)				✓
Não-Renováveis ($h - 48$)				✓
Não-Renováveis ($h - 168$)				✓

De forma a encontrar o melhor modelo para previsão do consumo de energia elétrica, também se fará uso dos modelos aleatório e de semanas misturadas, de forma a classificar as séries tendo em conta porções de todas as fases do ano (ver Secção 4.3.2). De referir que, neste caso, como os valores a estudar nunca são nulos, nem perto de zero, utiliza-se o indicador *MAPE* (ver Secção 3.2.3), para melhor perceção do peso das entradas. Importa referir que, contrariamente ao que se realiza na análise do preço da eletricidade, aqui atribui-se um maior peso modelo de semanas misturadas. Ainda que se prefira o modelo semanal ao aleatório, caso a diferença seja insignificativa, e como se trata de modelos seletivos (i.e., podem não seleccionar os dias, como os feriados, etc.), também se atribui prioridade à robustez do modelo (i.e. modelos com mais entradas).

5.1.2.9 Simulações para escolha do melhor modelo para 2016

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2016.

Tabela 5.18: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2016 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	3.85 %	3.84 %	4.14 %	4.21 %	3.71 %	3.49 %	2.84 %	2.54 %
MAE	219.45	219.33	232.92	237.15	214.67	199.04	159.99	143.04
MAPE'	3.90 %	3.90 %	4.13 %	4.20 %	3.81 %	3.54 %	2.84 %	2.53 %
RMSE	327.46	320.10	347.60	348.54	321.01	292.61	240.48	214.14
Tempo	2.14s	2.98s	2.18s	3.04s	4.95s	41.77s	5.82s	36.02s
Varição	-	-	-	-	↓ 0.14 %	↓ 0.35 %	↓ 1.30 %	↓ 1.67 %

Tabela 5.19: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2016 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	3.38 %	3.19 %	2.42 %	2.12 %	3.52 %	3.29 %	2.27 %	1.77 %
MAE	196.10	182.38	136.33	119.94	203.66	187.78	124.78	98.27
MAPE'	3.48 %	3.24 %	2.43 %	2.14 %	3.62 %	3.34 %	2.22 %	1.75 %
RMSE	278.79	264.30	192.56	172.69	292.33	262.91	166.19	133.63
Tempo	4.94s	1m 26s	8.16s	1m 20s	11.36s	4m 30s	17.35s	3m 25s
Varição	↓ 0.47 %	↓ 0.65 %	↓ 1.72 %	↓ 2.09 %	↓ 0.33 %	↓ 0.55 %	↓ 1.87 %	↓ 2.44 %

Analisando as Tabelas 5.18 e 5.19, percebe-se facilmente que a adição das variáveis cíclicas (Modelo A) e dos feriados (Modelo B) é bastante importante para o aumento de desempenho. Ainda assim, a adição de mais variáveis de entrada, já no Modelo C, quando comparado com o anterior, apresenta um decréscimo na performance, pelo que o modelo B é aquele a utilizar.

5.1.2.10 Simulações para escolha do melhor modelo para 2015

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2015.

Tabela 5.20: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2015 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	2.70 %	2.72 %	3.41 %	3.49 %	2.30 %	2.31 %	2.47 %	2.34 %
MAE	152.63	153.88	191.71	197.03	135.29	134.06	138.27	130.92
MAPE'	2.73 %	2.75 %	3.43 %	3.53 %	2.42 %	2.40 %	2.46 %	2.33 %
RMSE	204.65	2.75	299.00	302.67	208.79	183.05	227.09	206.28
Tempo	1.95s	1.99s	2.07s	2.31s	4.46s	43.45s	4.79s	31.17s
Variação	-	-	-	-	↓ 0.40 %	↓ 0.41 %	↓ 0.94 %	↓ 1.15 %

Tabela 5.21: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2015 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	2.45 %	2.28 %	2.17 %	1.91 %	3.61 %	2.43 %	2.27 %	1.67 %
MAE	143.26	131.46	121.99	107.00	206.14	140.61	125.36	92.81
MAPE'	2.56 %	2.35 %	2.18 %	1.91 %	3.69 %	2.52 %	2.25 %	1.66 %
RMSE	224.75	191.52	176.12	149.91	289.27	190.77	167.91	128.71
Tempo	5.09s	2m 29s	5.77s	1m 39s	9.92s	10m 28s	11.01s	10m 10s
Variação	↓ 0.25 %	↓ 0.44 %	↓ 1.24 %	↓ 1.58 %	↑ 0.91 %	↓ 0.29 %	↓ 1.14 %	↓ 1.82 %

No ano de 2015, como se verifica pelas Tabelas 5.20 e 5.21, a previsão reage positivamente à introdução das variáveis cíclicas e dos feriados (ainda que com uma ligeira subida no modelo semanal). De notar a perda de desempenho com a introdução das variáveis externas (Modelo C). Assim, sublinhando a leve perda de desempenho no modelo B, acoplada ao muito leve aumento nas restantes variantes, opta-se por escolher o modelo A como aquele a utilizar para o ano de 2015.

5.1.2.11 Simulações para escolha do melhor modelo para 2014

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2014.

Tabela 5.22: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2014 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	3.61 %	3.68 %	3.58 %	3.60 %	2.88 %	2.71 %	2.54 %	2.28 %
MAE	188.49	193.35	197.03	198.60	154.57	146.13	141.16	127.41
MAPE'	3.41 %	3.50 %	3.54 %	3.56 %	2.80 %	2.64 %	2.52 %	2.28 %
RMSE	253.61	251.79	297.66	296.67	223.44	214.96	209.18	189.48
Tempo	2.1s	2.22s	2.54s	2.14s	4.73s	48.28s	5.98s	39.88s
Varição	-	-	-	-	↓ 0.73 %	↓ 0.97 %	↓ 1.04 %	↓ 1.32 %

Tabela 5.23: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2014 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	2.96 %	2.61 %	2.22 %	1.93 %	3.84 %	3.30 %	2.37 %	1.77 %
MAE	156.56	140.49	123.61	107.99	208.48	176.29	129.93	98.00
MAPE'	2.83 %	2.54 %	2.22 %	1.94 %	3.77 %	3.19 %	2.33 %	1.76 %
RMSE	217.75	199.37	172.68	148.96	301.14	240.66	176.42	143.29
Tempo	5.56s	1m 45s	5.69s	1m 26s	10.09s	7m 46s	11.5s	4m 42s
Varição	↓ 0.65 %	↓ 1.07 %	↓ 1.36 %	↓ 1.67 %	↑ 0.23 %	↓ 0.38 %	↓ 1.21 %	↓ 1.82 %

Partindo dos valores apresentados nas Tabelas 5.22 e 5.23, nota-se mais uma vez o forte impacto das variáveis cíclicas e dos feriados. Mais ainda, realça-se uma muito pequena perda de desempenho no modelo B, mas coberta pelos avanços das outras variantes e pelo aumento de robustez, quando comparado com o modelo A. A introdução das variáveis externas, no modelo C, volta a introduzir erro, pelo que não será considerada. Ora, para o ano de 2014 também se estabelece o modelo B como aquele a utilizar para previsão via redes neuronais.

5.1.2.12 Simulações para escolha do melhor modelo para 2013

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2013.

Tabela 5.24: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2013 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	3.28 %	3.28 %	3.69 %	3.75 %	2.87 %	2.70 %	2.52 %	2.35 %
MAE	183.63	183.45	204.79	208.03	163.52	153.22	141.82	132.31
MAPE'	3.28 %	3.28 %	3.66 %	3.71 %	2.92 %	2.74 %	2.53 %	2.36 %
RMSE	244.69	245.35	315.66	319.32	222.57	213.95	215.03	197.13
Tempo	2.66s	2.23s	2.32s	2.46s	4.99s	43.2s	6.11s	39.35s
Variação	-	-	-	-	↓ 0.41 %	↓ 0.58 %	↓ 1.17 %	↓ 1.40 %

Tabela 5.25: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2013 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	2.94 %	2.71 %	2.33 %	2.15 %	3.90 %	2.87 %	2.25 %	1.78 %
MAE	166.34	154.18	130.22	120.99	218.97	161.11	124.38	98.44
MAPE'	2.97 %	2.75 %	2.32 %	2.16 %	3.91 %	2.88 %	2.23 %	1.77 %
RMSE	228.87	218.04	188.76	172.82	299.97	209.71	166.80	137.67
Tempo	5.99s	2m 25s	5.9s	1m 04s	9.25s	5m 55s	11.61s	5m 58s
Variação	↓ 0.34 %	↓ 0.57 %	↓ 1.36 %	↓ 1.60 %	↑ 0.62 %	↓ 0.41 %	↓ 1.44 %	↓ 1.97 %

No ano de 2013, ainda que visualmente pior do que os anteriores (ver Tabelas 5.24 e 5.25), também se estabelece o modelo B como aquele a utilizar. Note-se que, mesmo com um decréscimo praticamente inexistente (inferior a uma décima de um ponto percentual), atribui-se prioridade à robustez do modelo, associada ao aumento de desempenho da seleção aleatória. Perceba-se mais uma vez que com a seleção semanal e aleatória, podem ter sido selecionados períodos com mais ou menos feriados para um dos conjuntos de treino ou teste, resultando numa praticamente insignificante perda de performance.

5.1.2.13 Simulações para escolha do melhor modelo para 2012

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2012.

Tabela 5.26: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2012 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	4.00 %	4.01 %	4.01 %	4.06 %	3.02 %	2.81 %	2.70 %	2.46 %
MAE	228.82	229.48	220.26	224.43	176.26	162.07	150.81	137.28
MAPE'	4.07 %	4.08 %	3.95 %	4.03 %	3.14 %	2.88 %	2.70 %	2.46 %
RMSE	310.67	309.19	334.20	340.06	245.50	210.55	235.82	210.49
Tempo	2.57s	4.48s	2.79s	2.71s	6.29s	27.36s	5.92s	37.96s
Varição	-	-	-	-	↓ 0.98 %	↓ 1.20 %	↓ 1.31 %	↓ 1.60 %

Tabela 5.27: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2012 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	GRNN	ANN	GRNN	ANN	GRNN	ANN	GRNN	ANN
MAPE	2.96 %	3.02 %	2.36 %	2.03 %	4.30 %	3.13 %	2.46 %	1.65 %
MAE	171.54	172.28	132.35	114.47	242.78	180.10	134.44	90.12
MAPE'	3.05 %	3.06 %	2.37 %	2.05 %	4.32 %	3.20 %	2.42 %	1.62 %
RMSE	236.95	235.36	189.65	162.02	338.26	238.33	189.52	131.36
Tempo	6.89s	2m 48s	7.33s	1m 07s	12.77s	8m 40s	13.98s	3m 22s
Varição	↓ 1.04 %	↓ 0.99 %	↓ 1.65 %	↓ 2.03 %	↑ 0.30 %	↓ 0.88 %	↓ 1.55 %	↓ 2.41 %

Analisando as Tabelas 5.26 e 5.27, referentes aos dados de 2012, percebe-se novamente a importância das variáveis cíclicas e dos feriados. Ainda que seja possível obter uma leve redução de rendimento com a adição dos feriados, estabelece-se que a robustez do modelo deve ser tida em consideração pelo que, verificada a perda de rendimento no modelo C, se utiliza o modelo B para previsões finais.

5.1.2.14 Simulações para escolha do melhor modelo para 2011

Esta secção sintetiza os resultados das simulações realizadas com o objetivo de identificar o modelo mais adequado para previsão do consumo de energia elétrica para o ano de 2011.

Tabela 5.28: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos Base e A, relativos ao ano de 2011 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo Base				Modelo A			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	4.40 %	4.47 %	3.70 %	3.87 %	3.04 %	2.62 %	2.81 %	2.56 %
MAE	259.85	264.75	214.58	225.19	178.67	152.68	159.20	144.86
MAPE'	4.49 %	4.58 %	3.70 %	3.88 %	3.09 %	2.64 %	2.77 %	2.52 %
RMSE	387.67	387.85	332.07	338.46	262.23	223.49	247.98	231.34
Tempo	2.16s	3.13s	2.2s	2.14s	5.14s	32.15s	4.74s	24.23s
Variação	-	-	-	-	↓ 1.36 %	↓ 1.85 %	↓ 0.89 %	↓ 1.31 %

Tabela 5.29: Resultados dos testes de previsão do consumo de energia elétrica, por redes neurais, para os modelos B e C, relativos ao ano de 2011 (*MAE* e *RMSE* em *MWh*).

Modelo	Modelo B				Modelo C			
Seleção	Semanal		Aleatório		Semanal		Aleatório	
Rede	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
MAPE	2.72 %	2.48 %	2.40 %	2.05 %	3.96 %	3.01 %	2.48 %	1.69 %
MAE	161.15	146.86	137.92	137.92	231.28	178.53	140.55	96.98
MAPE'	2.79 %	2.54 %	2.40 %	2.40 %	4.00 %	3.09 %	2.43 %	1.67 %
RMSE	225.29	198.66	191.43	191.43	309.57	256.70	189.47	133.96
Tempo	5.69s	1m 10s	6.58s	54.43s	10.9s	8m 02s	10.23s	3m 25s
Variação	↓ 1.68 %	↓ 1.99 %	↓ 1.30 %	↓ 1.82 %	↓ 0.44 %	↓ 1.46 %	↓ 1.22 %	↓ 2.18 %

Sem grandes dúvidas neste caso, uma breve análise das Tabelas 5.18 e 5.19, indica facilmente que o modelo B será aquele a utilizar. Ora, com um aumento constante de desempenho com a adição das variáveis cíclicas e por conseguinte dos feriados, nota-se um aumento bastante razoável de desempenho, quando comparados com o modelo base. De notar também o decréscimo de desempenho do modelo C. Desta forma, também para o ano de 2011, o modelo a utilizar será o modelo B.

5.1.2.15 Resumo das simulações

Analisados de forma individual os anos em estudo, percebe-se rapidamente que, ainda que levemente afetados pela adição de variáveis (o erro base já é bastante reduzido), a robustez do modelo deve ser considerada. Perceba-se que, como para as previsões finais se faz uso do método adaptativo diário e não de métodos aleatórios ou semanais (não utilizáveis para previsões reais), a componente da robustez deve ser fortemente pesada, daí, a escolha dos modelos também ter sido moldada tendo em consideração este fator.

Tabela 5.30: Resumo dos modelos a utilizar para a previsão do consumo de energia elétrica, por ano.

Ano	2011	2012	2013	2014	2015	2016
Modelo	B	B	B	B	B	B

5.2 Previsões Finais

Nesta secção, apresentam-se os resultados finais das previsões do preço da eletricidade e do consumo de energia elétrica. De forma a aproximar o máximo possível os resultados a uma previsão real, estabelece-se que a previsão mais completa será aquela realizada pelo método de seleção adaptativo (dia-a-dia). Ainda assim, e de forma a comparar o desempenho, também se junta os resultados das previsões pelo método de seleção ordenada. De forma a coordenar uma melhor percepção dos resultados obtidos, clarifique-se que, numa primeira fase, se comparam os desempenhos dos dois métodos de amostragem de dados, seguidos duma comparação final entre metodologias de previsão.

Note-se que, assim como para todos os testes a realizar ao longo do exercício, todas as simulações são executadas com pesos de 75% e 25% para o conjunto de treino e teste, respetivamente. Neste caso, dada a utilização das técnicas de seleção ordenada e adaptativa, os pesos de 25% são sempre referentes ao último quarto do ano, notoriamente analisado como uma das secções com mais flutuação (no caso do preço da eletricidade).

Escolhidos os modelos a utilizar, e facilitada a tarefa de previsão pelo grande auxílio da aplicação desenvolvida, apresentam-se então os valores dos indicadores de desempenho para cada ano estudado, juntamente com um exemplo da variação da curva dos valores previstos, de um ano, no caso, 2016.

5.2.1 Previsões Finais do Preço da Eletricidade

Tabela 5.31: Resultados comparativos da previsão do preço da eletricidade, entre métodos de seleção ordenado e adaptativo, para os anos de 2011 a 2016.

Seleção		Ordenada		Adaptativa		Diferença	
Rede		GRNN	ANN	GRNN	ANN	GRNN	ANN
2011	MAE	7.28 €	8.37 €	6.59 €	6.45 €	↓ 9.48 %	↓ 22.94 %
	MAPE'	13.74 %	15.80 %	12.44 %	12.17 %	↓ 9.46 %	↓ 22.97 %
	RMSE	10.44 €	10.98 €	10.22 €	9.03 €	↓ 2.11 %	↓ 17.76 %
	Tempo	7.78s	1m 16s	10m 02s	2h 26m 51s	↑ 76 ×	↑ 115 ×
2012	MAE	10.94 €	14.04 €	9.79 €	8.51 €	↓ 10.51 %	↓ 39.39 %
	MAPE'	25.30 %	32.46 %	22.63 %	19.67 %	↓ 10.55 %	↓ 39.40 %
	RMSE	14.94 €	17.81 €	13.41 €	11.47 €	↓ 10.24 %	↓ 35.60 %
	Tempo	7.38s	1m 04s	9m 57	2h 58m 07s	↑ 80 ×	↑ 167 ×
2013	MAE	16.57 €	15.29 €	12.42 €	11.54 €	↓ 25.05 %	↓ 24.53 %
	MAPE'	31.72 %	29.27 %	23.78 %	22.10 %	↓ 25.03 %	↓ 24.50 %
	RMSE	21.47 €	20.08 €	17.67 €	16.21 €	↓ 17.70 %	↓ 19.27 %
	Tempo	6.42s	1m 26s	9m 16s	2h 35m 11s	↑ 86 ×	↑ 107 ×
2014	MAE	13.40 €	10.92 €	9.92 €	8.23 €	↓ 25.97 %	↓ 24.63 %
	MAPE'	26.96 %	21.98 %	19.96 %	16.56 %	↓ 25.96 %	↓ 24.66 %
	RMSE	17.48 €	13.94 €	13.62 €	10.69 €	↓ 22.08 %	↓ 23.31 %
	Tempo	6.16s	1m 33s	8m 45s	3h 10m 32s	↑ 84 ×	↑ 122 ×
2015	MAE	7.59 €	6.64 €	6.67 €	5.78 €	↓ 12.12 %	↓ 12.95 %
	MAPE'	14.79 %	12.95 %	13.01 %	11.27 %	↓ 12.04 %	↓ 12.97 %
	RMSE	10.03 €	8.55 €	9.06 €	7.56 €	↓ 9.67 %	↓ 11.58 %
	Tempo	7.07s	1m 42s	12m 19s	2h 56m 07s	↑ 104 ×	↑ 102 ×
2016	MAE	12.57 €	6.84 €	5.57 €	4.74 €	↓ 55.69 %	↓ 30.70 %
	MAPE'	22.15 %	12.05 %	9.82 %	8.36 %	↓ 55.67 %	↓ 30.62 %
	RMSE	14.52 €	8.24 €	7.28 €	6.08 €	↓ 49.86 %	↓ 26.21 %
	Tempo	6.13s	1m 38s	9m 32s	3h 05m 09s	↑ 92 ×	↑ 113 ×

Como é possível verificar pela Tabela 5.31, relativamente ao preço da eletricidade, a transição de um modelo de previsão em bloco (e.g., seleção ordenada) para um modelo de previsão adaptativo (e.g., adaptativo dia-a-dia), implica um grande aumento no desempenho, tanto das GRNN como das ANN. Entenda-se que, calculando a média das diferenças de todos os anos em estudo,

aumentou-se a performance da previsão em 23.14% e em 25.86% para as previsões por *GRNN* e por *ANN*, respetivamente, ainda que sujeitas a tempos de processamento muito superiores, no caso, 87 e 121 vezes acima, para as *GRNN* e para as *ANN*, respetivamente.

Tabela 5.32: Resultados finais e comparativos da previsão do preço da eletricidade, entre *GRNN* e *ANN*, para os anos de 2011 a 2016.

Seleção		Adaptativa		Diferença
Rede		<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>
2011	MAE	6.59 €	6.45 €	↑ 2.12 %
	MAPE'	12.44 %	12.17 %	↑ 2.17 %
	RMSE	10.22 €	9.03 €	↑ 11.64 %
	Tempo	10m 02s	2h 26m 51s	↓ 14 ×
2012	MAE	9.79 €	8.51 €	↑ 13.07 %
	MAPE'	22.63 %	19.67 %	↑ 13.08 %
	RMSE	13.41 €	11.47 €	↑ 14.47 %
	Tempo	9m 57	2h 58m 07s	↓ 17 ×
2013	MAE	12.42 €	11.54 €	↑ 7.09 %
	MAPE'	23.78 %	22.10 %	↑ 7.06 %
	RMSE	17.67 €	16.21 €	↑ 8.26 %
	Tempo	9m 16s	2h 35m 11s	↓ 16 ×
2014	MAE	9.92 €	8.23 €	↑ 17.04 %
	MAPE'	19.96 %	16.56 %	↑ 17.03 %
	RMSE	13.62 €	10.69 €	↑ 21.51 %
	Tempo	8m 45s	3h 10m 32s	↓ 21 ×
2015	MAE	6.67 €	5.78 €	↑ 13.34 %
	MAPE'	13.01 %	11.27 %	↑ 13.37 %
	RMSE	9.06 €	7.56 €	↑ 16.56 %
	Tempo	12m 19s	2h 56m 07s	↓ 13 ×
2016	MAE	5.57 €	4.74 €	↑ 14.90 %
	MAPE'	9.82 %	8.36 %	↑ 14.87 %
	RMSE	7.28 €	6.08 €	↑ 16.48 %
	Tempo	9m 32s	3h 05m 09s	↓ 18 ×

Comparando o desempenho das previsões obtidas por *GRNN* e por *ANN* (ver Tabela 5.32), rapidamente se entende que, as *ANN* apresentam uma performance relativamente superior, mas sem-

pre associadas a um tempo de processamento bastante superior, quando comparadas com aquele despendido pelas *GRNN*. Neste caso, para uma redução média de 11.26% do valor do *MAE*, as previsões tomam aproximadamente 17 vezes mais tempo a serem realizadas (valor médio).

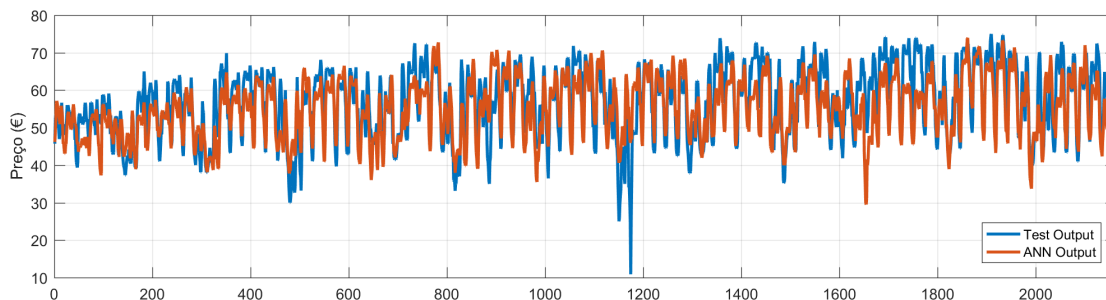


Figura 5.27: Resultados finais da previsão do preço em 2016 via *ANN*, por seleção autoadaptativa.

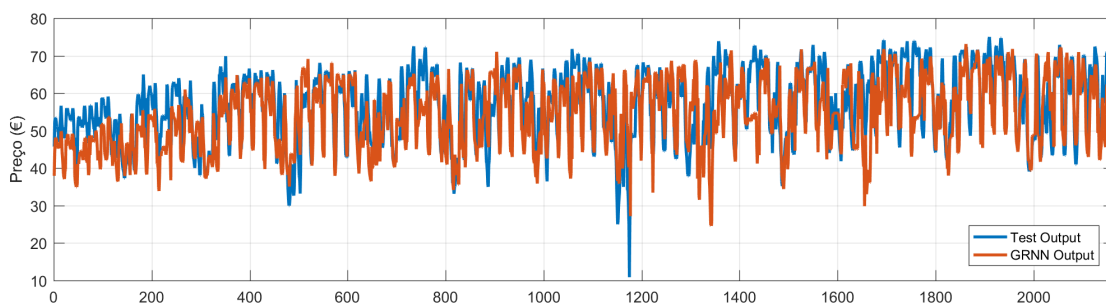


Figura 5.28: Resultados finais da previsão do preço em 2016 via *GRNN*, por seleção autoadaptativa.

Tomando como exemplo o ano de 2016, e relativamente ao preço da eletricidade, é possível verificar graficamente pelas Figuras 5.27 e 5.28 as curvas comparativas entre os valores reais e previstos por *GRNN* e *ANN*, respetivamente.

Neste caso, como os desempenhos são mais uma vez similares, é difícil identificar diferenças gráficas muito perceptíveis. Ainda que com uma diferença em valor absoluto de sensivelmente 11%, se verificarem certos pontos, onde as *ANN* garantem uma previsão um pouco mais precisa. De notar ainda um forte *outlier* perto do ponto 1200.

5.2.2 Previsões Finais do Consumo de Energia Elétrica

Tabela 5.33: Resultados comparativos da previsão do consumo de energia elétrica, entre métodos de seleção ordenado e adaptativo, para os anos de 2011 a 2016 (*MAE* e *RMSE* em *MWh*).

Seleção		Ordenada		Adaptativa		Diferença	
Rede		<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>	<i>ANN</i>
2011	MAPE	4.44 %	5.09 %	3.94 %	2.77 %	↓ 11.26 %	↓ 45.58 %
	MAE	250.21	287.60	218.76	155.59	↓ 12.57 %	↓ 45.90 %
	MAPE'	4.38 %	5.04 %	3.83 %	2.72 %	↓ 12.56 %	↓ 46.03 %
	RMSE	350.67	357.61	310.61	212.86	↓ 11.42 %	↓ 40.48 %
	Tempo	5.54s	29.31s	9m 25s	1h 25m 16s	↑ 101 ×	↑ 174 ×
2012	MAPE	4.16 %	4.99 %	3.92 %	3.50 %	↓ 5.77 %	↓ 29.86 %
	MAE	230.31	469.32	220.40	190.52	↓ 4.30 %	↓ 59.41 %
	MAPE'	4.11 %	4.81 %	3.93 %	3.40 %	↓ 4.38 %	↓ 29.31 %
	RMSE	362.26	482.25	347.02	291.46	↓ 4.21 %	↓ 39.56 %
	Tempo	8.25s	42.63s	7m 33s	1h 02m 42s	↑ 54 ×	↑ 87 ×
2013	MAPE	4.08 %	5.15 %	3.53 %	3.00 %	↓ 13.48 %	↓ 41.75 %
	MAE	233.84	290.56	201.25	166.13	↓ 13.94 %	↓ 42.82 %
	MAPE'	4.09 %	5.08 %	3.52 %	2.91 %	↓ 13.94 %	↓ 42.72 %
	RMSE	357.99	445.31	324.40	275.02	↓ 9.38 %	↓ 38.24 %
	Tempo	4.91s	34.17s	7m 23s	1h 04m 41s	↑ 89 ×	↑ 113 ×
2014	MAPE	3.96 %	4.59 %	3.55 %	3.12 %	↓ 10.35 %	↓ 32.03 %
	MAE	223.47	250.28	199.22	171.31	↓ 10.85 %	↓ 31.55 %
	MAPE'	3.96 %	4.43 %	3.53 %	3.03 %	↓ 10.86 %	↓ 31.60 %
	RMSE	363.34	404.54	319.04	284.27	↓ 12.19 %	↓ 29.73 %
	Tempo	4.87s	36.42s	8m 22s	1h 50m 33s	↑ 102 ×	↑ 181 ×
2015	MAPE	4.77 %	4.84 %	3.41 %	3.10 %	↓ 28.51 %	↓ 35.95 %
	MAE	254.73	264.21	182.92	165.05	↓ 28.19 %	↓ 37.53 %
	MAPE'	4.63 %	4.80 %	3.33 %	3.00 %	↓ 28.08 %	↓ 37.50 %
	RMSE	415.43	403.53	318.11	293.15	↓ 23.43 %	↓ 27.35 %
	Tempo	4.34s	14.67s	6m 08s	33m 25s	↑ 84 ×	↑ 136 ×
2016	MAPE	4.08 %	4.16 %	3.65 %	2.84 %	↓ 10.54 %	↓ 31.73 %
	MAE	229.21	233.83	207.06	159.21	↓ 9.66 %	↓ 31.91 %
	MAPE'	4.09 %	4.17 %	3.69 %	2.84 %	↓ 9.78 %	↓ 31.89 %
	RMSE	328.78	320.12	292.62	222.09	↓ 11.00 %	↓ 30.62 %
	Tempo	4.96s	24.42s	9m 19s	1h 13m 33s	↑ 112 ×	↑ 180 ×

Relativamente às diferenças entre os métodos de seleção ordenada e adaptativa, para a previsão do consumo de energia elétrica, através da análise da Tabela 5.33, entende-se que o modelo adaptativo se sobressai. Aqui, alcançam-se aumentos médios no desempenho de 13.32% para as *GRNN* e de 36.15% para as *ANN*, mas com um enorme aumento nos tempos médios de processamento, no caso, 90 e 145 vezes, respetivamente, quando comparados com os obtidos pela seleção ordenada.

Tabela 5.34: Resultados finais e comparativos da previsão do consumo de energia elétrica, entre *GRNN* e *ANN*, para os anos de 2011 a 2016 (*MAE* e *RMSE* em *MWh*).

Seleção		Adaptativa		Diferença
Rede		<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>
2011	MAPE	3.94 %	2.77 %	↑ 29.70 %
	MAE	218.76	155.59	↑ 28.88 %
	MAPE'	3.83 %	2.72 %	↑ 28.98 %
	RMSE	310.61	212.86	↑ 31.47 %
	Tempo	9m 25s	1h 25m 16s	↓ 8 ×
2012	MAPE	3.92 %	3.50 %	↑ 10.71 %
	MAE	220.40	190.52	↑ 13.56 %
	MAPE'	3.93 %	3.40 %	↑ 13.49 %
	RMSE	347.02	291.46	↑ 16.01 %
	Tempo	7m 33s	1h 02m 42s	↓ 7 ×
2013	MAPE	3.53 %	3.00 %	↑ 15.01 %
	MAE	201.25	166.13	↑ 17.45 %
	MAPE'	3.52 %	2.91 %	↑ 17.33 %
	RMSE	324.40	275.02	↑ 15.22 %
	Tempo	7m 23s	1h 04m 41s	↓ 8 ×
2014	MAPE	3.55 %	3.12 %	↑ 12.11 %
	MAE	199.22	171.31	↑ 14.01 %
	MAPE'	3.53 %	3.03 %	↑ 14.16 %
	RMSE	319.04	284.27	↑ 10.90 %
	Tempo	8m 22s	1h 50m 33s	↓ 12 ×
2015	MAPE	3.41 %	3.10 %	↑ 9.09 %
	MAE	182.92	165.05	↑ 9.77 %
	MAPE'	3.33 %	3.00 %	↑ 9.91 %
	RMSE	318.11	293.15	↑ 7.85 %
	Tempo	6m 08s	33m 25s	↓ 4 ×
2016	MAPE	3.65 %	2.84 %	↑ 22.19 %
	MAE	207.06	159.21	↑ 23.11 %
	MAPE'	3.69 %	2.84 %	↑ 23.04 %
	RMSE	292.62	222.09	↑ 24.10 %
	Tempo	9m 19s	1h 13m 33s	↓ 7 ×

Verificando a Tabela 5.34, rapidamente se percebe que as *ANN* obtêm um melhor desempenho do que as *GRNN*, mas associadas a tempos de treino bastante superiores. Neste caso, para a previsão do consumo de energia elétrica, por seleção adaptativa (dia-a-dia), ambas as variantes de redes neurais obtêm resultados excelentes, ainda que com uma ligeira vantagem para as *ANN*, com melhorias médias de apenas 16.47% quando comparando o *MAPE*, mas com tempos de execução médios 8 vezes superiores às *GRNN*.

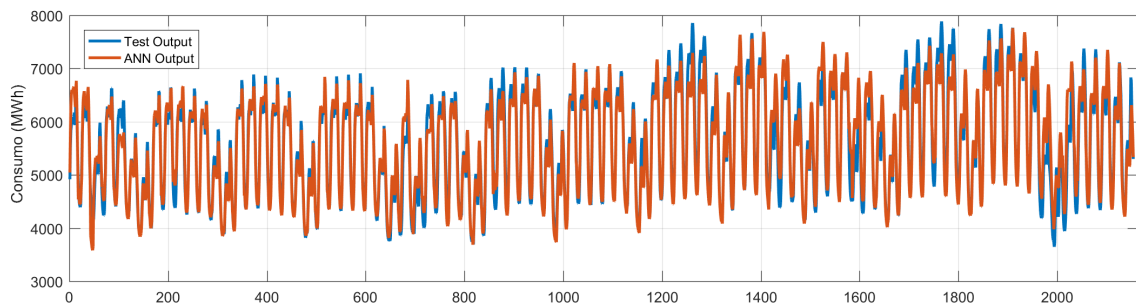


Figura 5.29: Resultados finais da previsão do consumo em 2016 via *ANN*, por seleção autoadaptativa.

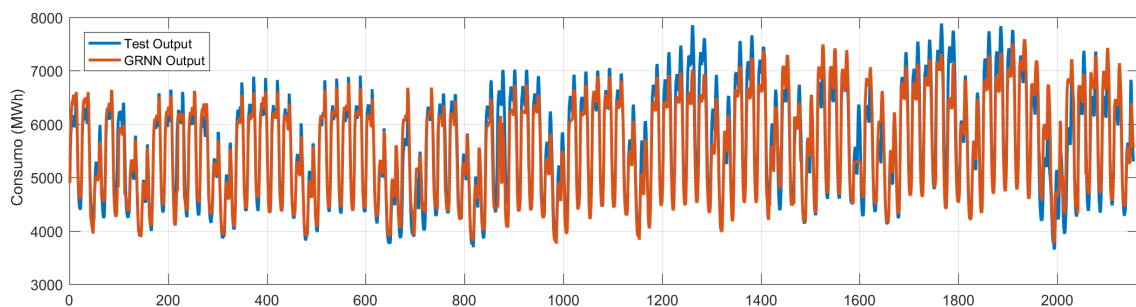


Figura 5.30: Resultados finais da previsão do consumo em 2016 via *GRNN*, por seleção autoadaptativa.

Tomando como exemplo o ano de 2016, relativamente ao consumo de energia elétrica, é possível verificar graficamente pelas Figuras 5.29 e 5.30 as curvas comparativas entre os valores reais e previstos por *ANN* e *GRNN*, respetivamente.

5.2.3 Previsão Experimental do Preço da Eletricidade com Seleção Horária

Ainda que distinta da previsão habitual para o dia seguinte, relativamente às *GRNN*, a previsão por seleção hora-a-hora apresenta bastantes melhorias, quando comparada com a seleção diária. Perceba-se que, com uma amostragem adaptativa hora-a-hora (ver Secção 4.3.2), o processo de previsão será deveras mais pesado. Pois bem, e como se trata apenas de uma previsão experimental, realiza-se apenas uma simulação, no caso, para a previsão dos últimos 25% do ano de 2016,

para o preço da eletricidade. Desta forma:

Tabela 5.35: Resultados comparativos da previsão experimental do preço da eletricidade, entre *GRNN* e *ANN*, por seleção adaptativa horária, para o ano de 2016.

Seleção		Adaptativa		Diferença
Rede		<i>GRNN</i>	<i>ANN</i>	<i>GRNN</i>
2016	MAE	2.69 €	3.38 €	↓ 25.65%
	MAPE'	4.73 %	5.96 %	↓ 26.00%
	RMSE	3.90 €	4.49 €	↓ 15.13%
	Tempo	3h 05m 09s	2d 15h 51m 21s	↓ 17 ×

Analisando a Tabela 5.35 facilmente se percebe que as *GRNN* seriam a escolha incontestável para previsão horária. Destaque o enorme aumento no tempo de processamento, na transição para as *ANN*, que tomaram praticamente 3 dias para a realização das previsões. No caso das *GRNN*, a simulação tomou apenas 3 horas, apresentando melhores resultados.

Perceba-se que a previsão horária é um processo muito moroso uma vez que realiza previsão de 25% de um ano, resultando na simulação de sensivelmente 2160 horas por ano, de forma individual. Para a previsão de cada hora, são tomados, em média, apenas 6s para as *GRNN* e 1m46s para as *ANN*.

Capítulo 6

Conclusões e Trabalho Futuro

6.1 Conclusões

Numa fase inicial da dissertação, as metas estipuladas passam pela análise de diferentes métodos de previsão do consumo de energia elétrica e do preço da eletricidade. Por conseguinte, após todo o trabalho realizado, serve o presente capítulo como forma de consolidação dos principais desfechos obtidos ao longo das diferentes fases do trabalho.

6.1.1 1ª Fase: Pré-processamento de Dados

Como primeira fase do exercício, procedeu-se à normalização dos dados obtidos para realização das previsões. Este conjunto de regras impostas, diferentes para os conjuntos de variáveis cíclicas ou não cíclicas, visa particularmente a gestão dos dados de forma a reduzir a redundância e a aumentar a sua integridade, assegurando-se um equilíbrio entre as escalas das séries de dados.

6.1.2 2ª Fase: Análise do Comportamento das Séries

Numa segunda etapa do trabalho, realiza-se uma profunda análise do comportamento do preço da eletricidade e do consumo de energia elétrica ao longo dos anos em estudo. Constata-se que o preço da eletricidade apresenta uma pertinácia característica no que toca a respeitar um padrão comum. Ora, é sensato afirmar que, excetuando frações de tempo singulares em cada um dos anos, o preço da eletricidade reivindica uma elasticidade capaz de transformar uma série de dados possivelmente equilibrada, numa outra ministrada de inúmeras flutuações, resultando numa superior incerteza para as previsões finais. Por comparação, o consumo de energia elétrica apresenta predicados firmes, rematando uma série de dados copiosamente uniforme, e preparada para fazer surgir resultados de previsão excelentes.

Daqui se distinguem também quais os atrasos mais congruentes para cada variável, mas que como provavelmente não utilizáveis numa previsão real, são mais uma vez importantes para a atribuição de prioridade ao prazo com o qual são executadas as previsões.

Note-se ainda que, pelo balanço realizado, entende-se que a secção em estudo, no caso, os últimos 25% de cada ano, são aquelas secções onde, geralmente, se germinam e apontam mais oscilações.

6.1.3 3ª Fase: Estudo e Escolha dos Métodos de Amostragem de Dados

Neste ponto do exercício, e precípua para uma boa interpretação das séries, estabelecem-se múltiplas formas de seleção de dados. Ora, depreende-se que os métodos de seleção aleatória e de semanas misturadas proporcionam interesse na seleção dos modelos de previsão, porque absorvem informação conexa com todas os segmentos da série. Dessa forma, e contrariamente aos demais métodos, que analisam exclusivamente a secção final de cada ano, estão aptos a executar um balanço mais integral e abrangente das séries anuais. Pois bem, estas duas técnicas são utilizadas para a seleção dos melhores modelos de previsão, com os restantes a serem empregues nas previsões finais. De salientar ainda o método adaptativo capaz de dispor de um comité de redes ao longo da série a prever, cada qual para um dia. Assim, administra de forma adaptativa, para a previsão de cada dia, a informação dos anteriores, de modo cíclico e sequencial.

6.1.4 4ª Fase: Identificação dos Modelos Finais para Previsão

Nesta etapa, fez-se uso das técnicas de amostragem de dados aleatória e de semanas misturadas, de forma a atestar estas previsões intermédias de dados de todo o ano. Ainda mais, acautelaram-se as previsões estabelecendo um modelo base que contemplava como entradas os deslocamentos da variável a prever com os atrasos de 48 e 168 horas. Daqui, relativamente ao preço da eletricidade, percebeu-se que a adição das variáveis cíclicas, dos feriados e das variáveis externas (i.e., previsão de carga e produções renováveis e não renováveis) é sempre favorável ao desempenho. Quanto à previsão do consumo de energia elétrica, e de uma forma geral, também corrobora a mais valia das variáveis cíclicas e dos feriados, sendo que a inclusão das variáveis externas (i.e., preço e produções renováveis e não renováveis) motivou uma perda de desempenho.

6.1.5 5ª Fase: Realização das Previsões

Num correto exercício de previsão, estabeleceu-se que esta tarefa será dirigida a uma hipótese o mais próxima possível de um cenário real. Desse jeito, utilizaram-se as técnicas de seleção ordenada e adaptativa dia-a-dia. Estabeleceu-se uma comparação entre estes mecanismos, onde se conclui que a seleção adaptativa, ainda que manifestamente mais morosa, triunfa de forma muito mais ilustre. Também aqui, se comparou finalmente a performance das *GRNN* e das *ANN*. Pois então, constatou-se que ambas as alternativas atingem resultados notáveis, ainda que com ligeira vantagem para as *ANN*. Não obstante da proximidade incontestável dos desempenhos, importa notar que os tempos de execução das *GRNN* são amplamente inferiores, pelo que se deve promover o uso das *GRNN* em exercícios com mais urgência.

6.1.6 Considerações Finais

Considerando um contexto real, a previsão torna-se um processo sujeito a análises, pelo que a identidade de cada caso deve ser explorada individualmente. Pois bem, o mistério associado aos resultados das previsões por ambos os métodos pode ser clarificado refletindo sobre o tipo de dados a estudar. Em virtude da necessidade de escolha dos modelos finais, estudaram-se as propriedades dos métodos de previsão para as seleções aleatórias e semanais. Este estudo permitiu esclarecer que, ainda que o trabalho tenha sido realizado em torno de apenas uma porção do ano, quando são simuladas diferentes secções, os resultados tendem a melhorar. Ainda mais, aprofunde-se a análise compreendendo que, ao contrário do normal, na previsão por seleção aleatória, as *GRNN* tendem a apresentar melhores resultados. Portanto, importa refletir sobre a possibilidade das *GRNN* realizarem mais efetivamente operações de interpolação, com as *ANN* a destacarem-se nos exercícios de extrapolação. A consulta de bibliografia sobre previsão mostra que o processo de amostragem aleatória é seguramente o mais comum. O trabalho desenvolvido nesta dissertação evidenciou claramente que esta opção conduz a desempenhos ilusoriamente otimistas. Esta é com certeza uma contribuição relevante deste trabalho.

Em jeito de conclusão, importa sublinhar que ambas as *GRNN* e as *ANN* se apresentam como soluções muito precisas. Porém, o primarismo subtil da arquitetura das *GRNN* a destaca-se pela positiva, originando um rendimento notável associado a tempos de execução muito discretos.

6.2 Aplicações Práticas da Previsão de Consumo e Preço

Associado ao constante aumento da cooperação para uma eficiente gestão da produção de energia elétrica (especificamente pelo contínuo aumento da produção de energias renováveis), a previsão do consumo de energia elétrica pode fornecer informação que sustente um aproveitamento otimizado destes diferentes tipos de produção. Nesse contexto, associado à possibilidade de escolha dos horizontes temporais na aplicação desenvolvida, caso seja necessário, a previsão poderá ser realizada de forma adaptativa, e de hora-a-hora. Ainda que para a previsão do preço da eletricidade, as diretrizes de mercado apontem para prazos bastante superiores (e daí a previsão ser feita com maior antecedência), no caso do consumo é exequível o planeamento de forma mais imediata. O carácter equilibrado do consumo permite claramente que a previsão horária possa ser ainda mais precisa. Já o preço, ainda que mais volátil, também poderia ser sujeito a uma pronta atuação das previsões, isto é, também se poderiam realizar previsões horárias, ainda que apenas aplicáveis ao mercado intradiário.

6.3 Dificuldades Encontradas

Tratando-se o tema da dissertação de uma abordagem inovadora de técnicas de previsão do consumo de energia elétrica e do preço da eletricidade, o cerne do trabalho vai de encontro às metodologias habituais da área da previsão. Dessa forma, não surgiram grandes dificuldades, com as existentes a passarem apenas por:

- Com a necessidade da obtenção dos dados para criação das redes e de respeitar as diretrizes de mercado, e como é habito na realização da previsão do preço da eletricidade e do consumo de energia elétrica, faz-se uso de dados meteorológicos médios nacionais como a temperatura do ar, força do vento, direção do vento, irradiância e pluviosidade. Neste exercício de previsão, tentou-se obter e mais tarde requereu-se os dados relativos aos anos em estudo, mas sem efeito. Ainda que sem a sua utilização os resultados tenham alcançado as expectativas, percebe-se que a adição destes dados poderia ter sido mais ou menos relevante;
- Dificuldade em obter dados relativos ao mercado ibérico, na sua totalidade.

6.4 Satisfação de Objetivos e Trabalhos Futuros

Em virtude dos objetivos estipulados numa fase inicial da dissertação, e considerando que o produto final do trabalho dispõe de diversos resultados (i.e., diferentes indicadores de desempenho) associados à previsão do consumo e do preço através de diversas metodologias e relativos a diversos anos, é possível afirmar que foi alcançada uma representatividade significativa. Ainda que se assuma que foram cumpridas as metas estabelecidas, fica, não obstante, uma nota para diferentes trabalhos futuros:

- Utilização de dados meteorológicos e dados ibéricos para criação de redes e consequente previsão;
- Utilização de uma base de dados com menor intervalo temporal (e.g. valores quarto-horários ao invés de horários);
- Seleção adaptativa com modelo atualizável (i.e., a adição de um novo dia implicava a remoção do dia/hora mais antigo;
- Agrupamento de dados (*Clustering*) por grau de semelhança;
- Escolha do parâmetro de suavização por metodologias diferentes (*Holdout Method* ou *The Wiggle-Method* [55]).

Referências

- [1] Edward A. Bender. *An Introduction to Mathematical Modeling*. John Wiley & Sons, Inc., 1978. Citado nas páginas 5, 6 e 7.
- [2] Bovas Abraham e Johannes Ledolter. *Statistical Methods for Forecasting*. John Wiley & Sons, Inc., 2005. Citado na página 8.
- [3] Chris Chatfield. *Time-Series Forecasting*. Chapman & Hall/CRC, 2001. Citado na página 9.
- [4] Jeffrey M. Woolridge. *Introductory Econometrics: A Modern Approach, Fifth Edition*. John Wiley & Sons, Inc., 2013. Citado na página 9.
- [5] REN. Centro de informação. Disponível em <http://www.centrodeinformacao.ren.pt/PT/InformacaoExploracao/Paginas/DiagramadeCargadaRNT.aspx>, acessado a última vez em 17 de Maio de 2017. Citado nas páginas 9, 30 e 38.
- [6] G. Udny Yule. Why do we Sometimes get Nonsense-Correlations between Time-Series? — A Study in Sampling and the Nature of Time-Series. *Journal of the Royal Statistical Society, Vol. 89, No. 1*, páginas 1–63, 1926. Disponível em <http://www.econ.uiuc.edu/~econ536/Papers/yule26.pdf>, acessado a última vez em 06 de Abril de 2017. Citado na página 9.
- [7] Eugen Slutsky. The Summation of Random Causes as the Source of Cyclic Processes. *Econometrica, Vol. 5, No. 2*, páginas 105–146, 1937. Disponível em <http://www.jstor.org/stable/1907241>, acessado a última vez em 06 de Abril de 2017. Citado na página 9.
- [8] Herman Wold. *A Study In The Analysis Of Stationary Time Series*. UPPSALA, 1938. Disponível em <https://archive.org/details/in.ernet.dli.2015.262214>, acessado a última vez em 06 de Abril de 2017. Citado nas páginas 9 e 10.
- [9] Spyros Makridakis e Michèle Hibon. ARMA Models And The Box Jenkins Methodology. *INSEAD Working Paper Series, Vol. 16, No. 3*, páginas 147–163, 1997. Disponível em [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1099-131X\(199705\)16:3%3C147::AID-FOR652%3E3.0.CO;2-X/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1099-131X(199705)16:3%3C147::AID-FOR652%3E3.0.CO;2-X/abstract), acessado a última vez em 17 de Maio de 2017. Citado na página 10.
- [10] George E. P. Box, Gwilym M. Jenkins, e Gregory C. Reinsel. *Time Series Analysis, Forecasting and Control*. John Wiley & Sons, Inc., 1970. Citado na página 10.
- [11] J. Phillip Cooper. Asymptotic Covariance Matrix of Procedures for Linear Regression in the Presence of First-Order Autoregressive Disturbances. *Econometrica, Vol. 40, No. 2*, páginas 305–310, 1972. Disponível em www.jstor.org/stable/1909408, acessado a última vez em 17 de Abril de 2017. Citado na página 10.

- [12] Stephen Beveridge e Charles R. Nelson. A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle. *Journal of Monetary Economics*, Vol. 7, No. 2. Disponível em <http://www.sciencedirect.com/science/article/pii/0304393281900404>, acessado a última vez em 17 de Abril de 2017. Citado na página 10.
- [13] Robert L. Winkler e Spyros Makridakis. The Combination of Forecasts. *Journal of the Royal Statistical Society. Series A (General)*, Vol. 146, No. 2, páginas 150–157, 1983. Disponível em <http://www.jstor.org/stable/2982011>, acessado a última vez em 21 de Abril de 2017. Citado na página 10.
- [14] B. S. Everitt e A. Skrondal. *The Cambridge Dictionary Of Statistics*. Cambridge University Press, 2010. Citado na página 11.
- [15] Theresa Hoang Diem Ngo. The Box-Jenkins Methodology for Time Series Models. *SAS Global Forum*, 2013. Paper 454-2013, disponível em <http://support.sas.com/resources/papers/proceedings13/454-2013.pdf>, acessado a última vez em 19 de Abril de 2017. Citado na página 12.
- [16] J. Nuno Fidalgo. Séries temporais. Conteúdos da U.C. de Técnicas para Previsão, disponível em https://sigarra.up.pt/feup/pt/conteudos_geral.ver?pct_pag_id=249640&pct_parametros=pv_ocorrencia_id=385675#58444, acessado a última vez em 05 de Abril de 2017. Citado na página 12.
- [17] Simon Haykin. *Neural Networks: A Comprehensive Foundation*. Pearson Education, 1999. Citado nas páginas 13, 14, 15, 20, 22, 23 e 26.
- [18] Warren S. McCulloch e Walter Pitts. A Logical Calculus Of The Ideas Immanent In Nervous Activity. *Bulletin of Mathematical Biophysics*, Vol. 5, No. 1/2, páginas 99–115, 1943. Disponível em <http://www.cs.cmu.edu/~epxing/Class/10715/reading/McCulloch.and.Pitts.pdf>, acessado a última vez em 12 de Maio de 2017. Citado nas páginas 13 e 19.
- [19] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, Vol. 65, No. 6, páginas 386–408, 2014. Citado na página 13.
- [20] M. Minsky e S. Papert. *Perceptrons*. MIT Press, 1969. Citado na página 14.
- [21] Marvin Minsky e Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry*. The MIT Press, 1987. Citado na página 14.
- [22] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 79, No. 8, páginas 2554–2558, 1982. Citado na página 14.
- [23] D. E. Rumelhart, G. E. Hinton, e R. J. Williams. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1. páginas 318–362, 1986. Citado na página 14.
- [24] P. J. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Tese de doutoramento, Harvard University, 1974. Citado na página 14.

- [25] Guoqiang Zhang, B. Eddy Patuwo, e Michael Y. Hu. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, Vol. 14, No. 1, páginas 35–62, 1998. Disponível em <http://www.sciencedirect.com/science/article/pii/S0169207097000447>, acessado a última vez em 12 de Maio de 2017. Citado na página 14.
- [26] Paulo Cortez e José Neves. Redes neuronais artificiais, 2000. Universidade do Minho. Citado nas páginas 15, 17 e 19.
- [27] Raúl Rojas. *Neural Networks: A Systematic Introduction*. Springer-Verlag, Berlin, 1996. Citado nas páginas 17 e 24.
- [28] Michael Jordan. Why the logistic function? A tutorial discussion on probabilities and neural networks. Relatório técnico, Massachusetts Institute of Technology, 1995. Disponível em https://www.ics.uci.edu/~dramanan/teaching/ics273a_winter08/homework/jordan_logistic.pdf, acessado a última vez em 12 de Maio de 2017. Citado na página 19.
- [29] Dan W. Patterson. *Artificial Neural Networks: Theory and Applications*. Prentice-Hall Series in Advanced Communications. Prentice Hall, 1996. Citado nas páginas 22 e 26.
- [30] D.O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. John Wiley & Sons, Inc., 1949. Citado na página 23.
- [31] D. S. Broomhead e David Lowe. Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, Vol. 2, páginas 321–355, 1988. Disponível em <http://sci2s.ugr.es/keel/pdf/algorithm/articulo/1988-Broomhead-CS.pdf>, acessado a última vez em 13 de Maio de 2017. Citado na página 24.
- [32] Martin Riedmiller. Advanced Supervised Learning in Multi-Layer Perceptrons - From Backpropagation to Adaptive Learning Algorithms, Vol. 16, No. 3. páginas 265–278, 1994. Disponível em <http://www.sciencedirect.com/science/article/pii/S0920548994900175>, acessado a última vez em 13 de Maio de 2017. Citado na página 24.
- [33] Vladimiro Miranda. Redes Neuronais - treino por retropropagação. 2007. Versão 1.1. Citado na página 24.
- [34] Kenneth Levenberg. A Method For The Solutions Of Certain Non-linear Problems in Least Squares. *Quarterly of Applied Mathematics*, Vol. 2, No. 2, páginas 164–168, 1944. Disponível em <https://www.jstor.org/stable/43633451>, acessado a última vez em 21 de Abril de 2017. Citado na página 25.
- [35] Donald W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, No. 2, páginas 431–441, 1963. Disponível em <http://epubs.siam.org/doi/10.1137/0111030>, acessado a última vez em 21 de Abril de 2017. Citado na página 25.
- [36] M. T. Hagan e M. B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, páginas 989–993, Nov 1994. Disponível em <http://ieeexplore.ieee.org/document/329697/>, acessado a última vez em 21 de Abril de 2017. Citado na página 25.

- [37] M.T. Hagan, H.B. Demuth, M.H. Beale, e Orlando De Jesus. *Neural Networks Design*. Martin Hagan, 2014. Citado na página 25.
- [38] OMIE. Mercado da eletricidade. Conteúdo do sítio da OMIE, disponível em <http://www.omie.es/pt/principal/mercados-e-produtos/conheca-o-nosso-mercado>, acedido a última vez em 03 de Maio de 2017. Citado nas páginas 28 e 29.
- [39] OMIE. Resultados do mercado. Conteúdo do sítio da OMIE, disponível em <http://www.omie.es/files/flash/ResultadosMercado.swf>, acedido a última vez em 03 de Maio de 2017. Citado nas páginas 29 e 31.
- [40] REN. O que fazemos: O setor elétrico. Conteúdo do sítio da REN, disponível em http://www.ren.pt/pt-PT/o_que_fazemos/eletricidade/o_setor_eletrico/, acedido a última vez em 03 de Maio de 2017. Citado na página 30.
- [41] Shaikh Abdul Hannan, R. R. Manza, e R. J. Ramteke. Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis. *International Journal of Computer Applications*, Vol. 7, No. 13, páginas 7–13, Oct 2010. Publicado por Foundation of Computer Science, disponível em <http://www.ijcaonline.org/archives/volume7/number13/1325-1799>, acedido a última vez em 04 de Maio de 2017. Citado nas páginas 33 e 34.
- [42] M.M. Tripathi, K.G. Upadhyay, e S.N. Singh. Short-Term Load Forecasting Using Generalized Regression and Probabilistic Neural Networks in the Electricity Market. *The Electricity Journal*, Vol. 21, No. 9, páginas 24–34, Nov 2008. Disponível em <http://www.sciencedirect.com/science/article/pii/S1040619008002030>, acedido a última vez em 04 de Maio de 2017. Citado na página 34.
- [43] Gaurav Kumar e Hasmat Malik. Generalized Regression Neural Network Based Wind Speed Prediction Model for Western Region of India. *Procedia Computer Science*, Vol. 93, páginas 26–32, 2016. Disponível em <http://www.sciencedirect.com/science/article/pii/S187705091631417X>, acedido a última vez em 04 de Maio de 2017. Citado na página 34.
- [44] MathWorks. Generalized regression neural networks. Conteúdo do sítio da MathWorks, disponível em <https://www.mathworks.com/help/nnet/ug/generalized-regression-neural-networks.html>, acedido a última vez em 16 de Maio de 2017. Citado na página 34.
- [45] Donald F. Specht. A General Regression Neural Network. *IEEE Transactions on Neural Networks*, Vol. 2, No. 6, páginas 568–576, Nov 1991. Citado nas páginas 34 e 36.
- [46] Felix Heimes. The Normalized Radial Basis Function Neural Network. *Systems, Man, and Cybernetics, 1998 IEEE International Conference*, Vol. 2, páginas 1609–1614, Oct 1998. Citado na página 35.
- [47] M. Zhong, D. Goggeshall, E. Ghaneie, T. Pope, M. Rivera, M. Georgiopoulos, G. Anagnostopoulos, M. Mollaghasemi, e S. Richie. Gap-Based Estimation: Choosing the Smoothing Parameters for Probabilistic and General Regression Neural Networks. *MIT Press Journals*, Vol. 19, No. 10, páginas 2840–2864, Oct 2007. Citado na página 36.

- [48] OMIE. Resultados de mercado. Disponível em <http://www.omie.es/files/flash/ResultadosMercado.swf>, acessado a última vez em 17 de Maio de 2017. Citado na página 38.
- [49] E. Kreyszig. *Advanced Engineering Mathematics*. John Wiley & Sons, 2010. Citado na página 39.
- [50] R.J. Hyndman e G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, 2014. Citado na página 40.
- [51] Ray C. Fair. *Evaluating the predictive accuracy of models*, volume 3 de *Handbook of Econometrics*. Elsevier, 1986. Citado nas páginas 40 e 41.
- [52] Zhuo Chen e Yuhong Yang. *Assessing Forecast Accuracy Measures*, 2004. Citado na página 40.
- [53] Elisabeth Woschnagg e Jana Cipan. *Evaluating Forecast Accuracy*, 2004. Citado na página 41.
- [54] The MathWorks Inc. *MATLAB version 9.1.0.441655 (R2016b)*. Natick, Massachusetts, 2016. Citado nas páginas 43, 44, 45, 47 e 52.
- [55] Matthias M. Bauer. *General Regression Neural Network, GRNN - A Neural Network for Technical Use*. University of Wisconsin-Madison, 1995. Disponível em <https://minds.wisconsin.edu/handle/1793/7779>, acessado a última vez em 10 de Junho de 2017. Citado na página 98.